



November 25, 1976

Half greater than whole

WITH the steady decline in the number of those who go into productive industry in Britain and the steady rise in the number of people worrying about the problem, before the end of the century we expect these numbers to be equal and sometime early in the twenty-first century to see the disappearance of the last industrial worker. The Commons Select Committee on Science and Technology is the latest to agonise over the trend (Commons paper 680, £1.15), and if anything may have helped to speed it on its way by publishing some bitter remarks from some of its witnesses. From Lord Bowden, formerly Principal of the University of Manchester Institute of Science and Technology, for example, about the hazards of industrial life, the rather unpleasant environment and the contempt of society at large for those who create the wealth which everyone wants to spend. Or from the Chairman of EMI, who declared that you wouldn't choose the industrial environment for security, money or professional satisfaction. Or even from Mr Varley, the Secretary of State for Industry, who said that 'to be a technologist or a manager within industry has been a pretty lousy job'.

The committee, in looking at university-industry relations, has bitten off more than it can reasonably chew. Some recommendations (see also page 311)—on the future of the National Research Development Corporation, on the moving of basic research out of governmental research laboratories, on the functioning of the Advisory Council on Applied Research and Development, on the establishment of a Minister of Science and Technology, and on the extension of the customer-contractor principle into the Science Research Council (SRC) with the Department of Industry (DoI) as customer for the applied research—seem too hastily put together and based on too scanty evidence to be taken seriously. For instance on this last issue the committee has failed to consider that SRC, unlike the other research councils, supports little applied research in its own institutions, and that the proposal would, therefore, let the DoI (already well equipped with laboratories) choose to support or kill research projects in university laboratories—not, surely, the committee's intention.

It would be unfortunate, however, if the veritable deluge of half-worked-out proposals in the second half of the report were allowed to overshadow a valuable analysis, in the first half, of the development of post-war governmental support for science and technology education. It discusses the failures of new institutions and institutions with new names to convince bright students in large numbers that industry offered exciting

prospects—a failure compounded by industry's own inability to project itself as a place where highly qualified scientists and technologists would reach the top. The report echoes a DoI assessment that 'the overall trend is a decline in the numbers and quality of qualified scientists and engineers and of supporting staff in key areas of manufacturing'.

Status and motivation seem to emerge as the villains of the piece. Young people do not see engineering as a worthy profession in which to exercise intellectual skills: as a result there is too little understanding of what industry in all its complexity is about, or willingness to participate wholeheartedly amongst those 'forced' into it. The Select Committee, however, having inveighed against 'the distressing habit of attempting to bestow status (on academic institutions) by changing names', proceeds to propose as one recommendation that Imperial College, UMIST, University of Strathclyde and some former Colleges of Advanced Technology should be labelled SISTERS (Special Institutions for Scientific and Technological Education and Research). This is a revival of a 1963 idea to give technology greater prominence. The Rector of Imperial College will no doubt be mildly surprised to find that the committee propose that applied science and engineering at Cambridge be also made into a SISTER 'to ensure the high status of SISTERS is recognised by the academic community'.

The SISTER idea, unfortunately, is as poorly worked out as most of the committee's other ideas. To issue these institutions with revised charters limiting their functions to training and research in engineering and applied sciences would be to block all the fertile, increasingly necessary and in Britain often poorly developed cross-linkages into pure science, mathematics and economics. A dull narrow-minded product you'd get.

And yet the committee do at times get near to the centre of the problem: from early on in education the claims of industry and engineering are nothing like clearly enough presented. How many schools ensure that their pupils see a cross-section of industrial environments? How many industries make a serious attempt to get through to undergraduates, even to the extent of using them in vacations? The claims of an academic life are constantly being impressed on youngsters by their educational environment. Very often their only sight of industry before graduating is of strike picket lines on television news. There are some cautious ventures going ahead in this direction of mutual understanding, but it is up to the DoI, the Department of Education and Science and the Confederation of British Industry to stimulate much more. □



Stocks and shares

Questions about fish stocks lie behind Europe's fishing limits debate. **D. H. Cushing** discusses the state of stocks in the North-East Atlantic

THE conservation of fish stocks in the North-East Atlantic (between Cape Farewell and Gibraltar) started with the Overfishing Convention of 1946 which prescribed minimum landing sizes for demersal fishes in the ports and minimum mesh size in the trawl cod ends at sea. The Convention was activated in the Permanent Commission in 1953 which itself was replaced by the North-East Atlantic Fisheries Commission (NEAFC) in 1964 in order to acquire the power to regulate herring stocks as well as demersal ones. Sixteen nations have adhered to it voluntarily to conserve fish stocks internationally. In the last three years the Third Law of the Sea Conference has established a quite different trend towards coastal state management.

The present object of conservation is the maximum sustained yield (MSY) in weight. Men dispute whether it is a proper objective and propose others of various forms. In the North-East Atlantic, however, most stocks are overexploited and fishing effort should be reduced to obtain the MSY. This means that fewer ships at less cost would gain a greater catch and the average catch of each vessel or catch per unit of effort would be considerably increased. In his book *The Fish Gate* (Faber, London; 1943), Michael Graham named this principle the Great Law of Fishing.

The major stocks taken by fishermen in the North-East Atlantic comprise the cod-like fishes (cod, haddock, whiting and saithe), the flatfish (plaice and sole) and the pelagic species (herring, sprat and mackerel). The stocks are estimated by natural regions such as the North Sea, Irish Sea, Barents Sea and so on, between which exchange of individuals is usually low. The state of the stocks by species is summarised in the accompanying table; they are given by region (such as the North Sea) and are classed in four groups according to the reduction in fishing effort needed to obtain the MSY.

There are three stocks at or approaching the MSY with present measures: North Sea plaice, North Sea whiting and the cod in the Barents Sea. A slight restraint is required in five stocks: West Scotland herring, Iceland saithe, Faroe haddock and West Scotland cod and haddock. Eight stocks would be brought to the MSY

with moderately reduced fishing effort: Faroe cod, saithe in the North Sea and in the North-East Arctic, and both flatfish species in the Irish Sea, Bristol Channel and the English Channel. Unfortunately, there are fourteen stocks in need of heavy reduction in fishing effort: cod in the North Sea, Irish Sea and Bristol Channel, haddock in the North-East Arctic, North Sea, Irish Sea and Bristol Channel, whiting in the Irish Sea and Bristol Channel, sole in the North Sea and Irish Sea and herring in the North Sea and Irish Sea. The state of sprat and mackerel stocks in the North Sea and of mackerel stocks in the Western Approaches is uncertain.

Three reasons

There are three reasons why, after twenty years, most fish stocks in the North-East Atlantic remain over-exploited. The first concerns a limit in mesh regulation. The impact of an increase in fishing effort on a single species can be lessened to some degree by enlarged meshes in the trawl cod ends, but in a mixed fishery the meshes are fitted to the smallest fish species and any increase in fishing effort on the larger animals cannot be accommodated in this way. In the North Sea, mesh regulation was designed mostly for soles and haddock but the benefit for larger species like cod and plaice was less. In the North Sea and elsewhere in the North Atlantic, minimum mesh sizes were agreed by the end of the

Sorry, for copyright reasons some images on this page may not be available online

Quota problems go deep in Denmark...

1950s. But during the 1960s a considerable increase in fishing effort occurred all throughout these regions. In the International Commission for North-West Atlantic Fisheries (ICNAF), quotas were introduced in 1970 and in NEAFC in 1974. From North-West Atlantic experience, quotas stabilise fishing and prevent further expansion.

The collapse of herring fisheries throughout the North-East Atlantic illustrates the second reason for over-exploitation. The Downs stock, exploited by the East Anglian fishery, failed in the late 1950s, followed by other stocks in the North Sea; when NEAFC took power to regulate herring fisheries it did not succeed in doing so. When the Norwegian fishery on the Atlanto-Scandian herring stock collapsed in 1967, no action was taken. A contributory but minor cause of failure was the mistaken belief that fishermen must take fewer herring than the whales.

State of the stocks in the North-East Atlantic: reduction in fishing needed to obtain MSY

	None	Slight	Moderate	Heavy
Cod	North-East Arctic ¹	W. Scotland	Faroe	North Sea Irish Sea Bristol Channel
Haddock		Faroe W. Scotland		North-East Arctic North Sea Irish Sea Bristol Channel
Whiting	North Sea			Irish Sea Bristol Channel
Saithe		Iceland	North-East Arctic North Sea	
Plaice	North Sea		Irish Sea Bristol Channel English Channel	
Sole			Bristol Channel English Channel	North Sea Irish Sea
Herring		W. Scotland		North Sea Irish Sea Celtic Sea

D. H. Cushing is Deputy Director of the Fisheries Laboratory, Lowestoft, UK.

¹At the present pattern of exploitation.

dogfish, cod and gannets that driftermen saw round their nets.

The major cause was the scientific dogma that fishing could never reduce the magnitude of the incoming year classes. Today we distinguish growth overfishing from recruitment overfishing. In growth overfishing the loss in numbers is greater than need be to give the maximum yield from the growth of individuals, and it is no accident that such individuals grow by an order of magnitude during their adult lives. In recruitment overfishing (often in herring-like fishes that grow little during their adult lives) the magnitude of the incoming year classes is reduced by fishing—an interference with breeding, as the fishermen used to say. The collapse of the herring stocks in the North-East Atlantic has been due to recruitment overfishing.

One of the difficulties was the high variability of recruitment; the dependence of recruitment on parent stock looked like a scatter diagram and sometimes the true relationships were masked. Indeed the belief that recruitment was independent of parent stock was mistakenly reinforced. The important conclusion, however, is that the overexploitation of the herring fisheries in the North-East Atlantic was less due to institutional failure than to the inability of scientists to understand the problems they faced. It remains true that the failure of NEAFC to take prompt and effective action was due in part to the difficulties of agreeing upon measures which would necessarily severely restrict the fisheries economies of some member states.

Considerable overexploitation

The development of industrial fisheries for fish meal in the North Sea has caused considerable overexploitation indirectly. During the 1960s herring were caught for fish meal on their spawning grounds and the trawlers were allowed to take 10% by-catch of those protected species which gathered to eat the herring spawn on the sea bed. Today more than two million tons of fish are taken from the North Sea for fish meal and the by-catches of protected species can be considerable. The sandeel fishery is "clean" and little else but sandeels is caught. In the fishery for Norway pout (which are very small gadoids) in the Northern North Sea, however, small haddock and whiting are caught in large numbers. In some years more than 100,000 tons of small haddock have been turned into fish meal, and in 1976 75,000 tons of small whiting were caught for the same purpose.

Norway pout are caught with small-meshed nets, legal under NEAFC rules, provided that the catch of under-sized protected species is less than 10%;

the total catch of Norway pout in 1975 exceeded 800,000 tons. The real difficulty is that fishermen exploiting adult gadoids for human consumption must not land fish below the minimum landing size and must discard them at sea, whereas fish meal fishermen can land large quantities of them; the large industrial fleets are seen everywhere in the North Sea and the traditional fishermen for human consumption see the small gadoids being taken which they themselves have to discard. Further, a rough calculation shows that 300,000 tons of Norway pout for fish meal are less valuable than 20,000 tons of haddock, had they been allowed to grow into the adult fishery.

After two decades of recruitment overfishing, the remaining herring stocks in the North Sea need drastic conservation measures. This need has at last been recognised by the industrial fishing nations who have agreed to a ban on all directed fishing for herring in the North Sea. But juvenile herring are caught in the industrial fishery for sprats; in 1975, about 600,000 tons of sprats were taken and the by-catch of small herring in the Danish catches was substantial. In other words, the industrial fishery for sprats not only prevents any chance of recovery of herring stocks, it also hastens the chance of extinction. A fair proportion of the world's catch of fish ends as fishmeal because of the high demand for it, and many industrial fisheries, like that for capelin in the Barents Sea or for anchoveta off Peru, exploit resources that cannot be used at all for human consumption. The problem peculiar to the North Sea is the conflict in use between fisheries for industrial purposes and those for protected species.

The three sources of overexploitation were recruitment overfishing, the conflict between fisheries for industrial species and those for protected species,

and the increase in fishing effort in ships, fishermen and gears. The herring and sole stocks in the North Sea are in danger from recruitment overfishing, but the problem has been recognised and steps are being taken in the scientific working groups to resolve it, even if the herring problem has not yet been solved in the Commission itself. The problems of the industrial fisheries are more difficult because of the large investment in ships and processing equipment in Denmark. Possible solutions include the diversion of such fleets on to stocks which would yield no by-catch of protected species, or to establish areas and seasons in the Norway pout and sprat fisheries when the by-catches of protected species are reduced to tolerable levels. The problem must be solved because it is hard to persuade the non-industrial fishermen to accept quotas or effort reduction when the industrial fleets are allowed to catch small gadoids and small herring.

The table shows that fishing effort should be extensively reduced throughout the North-East Atlantic. The introduction of quotas in 1975 has stabilised fishing effort; in ICNAF with a longer experience of quotas since 1970, an effort restraint was also introduced in 1976. The effort exerted on all the stocks on the right half of the table is much too great, which means that fleets in the North Atlantic will have to be reduced to some degree. There are few stocks left unexploited, but one is large, that of blue whiting off the west coast of the British Isles. Stocks are estimated at about 8–15 million tons and the potential catches could be large, but the exploitation has hardly started. Although some existing vessels could be used to catch the blue whiting, it remains the case that fishing on the great majority of other stocks in the North-East Atlantic must be reduced, in some cases drastically. □

Sorry, for copyright reasons some images on this page may not be available online

400 GeV . . . and beyond

Experiments begin soon on the 400-GeV Super Proton Synchrotron (SPS) at CERN, Geneva.

David Davies has been visiting the laboratory

SPS comes on line within budget and on schedule. After six years of awaiting the complicated multinational approval, and a further five of building an accelerator of radius 1.2 km which circles some 40 metres below ground, the physics starts in earnest in early December when the first high-energy particles are fed to a whole array of experiments. Major construction still continues, as the plans are for two completely separate experimental areas, only one of which is yet finished. But already questions are being asked: what do we build next?

CERN started life with a 600 MeV synchrocyclotron (SC) commissioned in 1957 and still in use. In 1959 a 28-GeV proton synchrotron (PS), fed successively by a 550-KeV electrostatic generator and a 50-MeV linear accelerator, was added. Although there is still an active physics programme on the PS, it also serves to feed the two other machines—the intersecting storage rings (ISR) and the SPS. ISR was completed in 1971, and comprises two rings of diameter 300 metres in which 28 GeV protons can travel for up to 40 hours in opposite directions crossing over and colliding at eight equally spaced positions—at six of which experiments are sited. The performance of the ISR in terms of storage current and luminosity is continually improving, and plans for its use well into the 1980s are being discussed at present (see page 314).

The advantage of machines in which beams travelling in opposite directions collide is that almost all the energy is available for particle production: SPEAR at Stanford and the projected PETRA at Hamburg are further examples, in which electrons and positrons are ranged against each other. In accelerators which bombard stationary targets, a large fraction of the energy goes in recoil of the target

particles. On the other hand, storage rings are restricted to a much smaller range of experiments—in other accelerators the primary beam is very often converted at a target to secondary beams of completely different particles.

The protons for the SPS are peeled off the PS in a continuous ribbon which eventually fills ten-elevenths of the SPS's seven-kilometre circumference. Once in the ring the 10^{13} protons encounter the magnetic field both of dipoles (which bend the protons in a horizontal plane) and quadrupoles (which keep the beam focused). There are 744 dipoles (peak field 1.4 Teslas) and 216 quadrupoles and the mean power consumption is 34 MW. The vacuum is 10^{-7} Torr. Acceleration of protons occurs at only one point on the ring, in a straight stretch of 40 metres where there are two radio-frequency cavities. Each journey through this section adds about 2.5 MeV to the protons' energy. After 150,000 circuits, which take 3.7 seconds, the protons are up to an energy of 400 GeV and ready to be ejected into either of the experimental areas, West or North. West Area is the first to be completed, North follows early in 1978.

The proton ribbon can either be extracted all in one turn or can be progressively peeled off over a period of up to a second. Bubble chamber experiments tend to need the short burst, electronic counter experiments the longer time. The beam can be used in the West Area in three ways:

- It can be directed on to an underground target at up to 40 GeV to produce pions and kaons which decay within 430 metres of their journey to the surface into muons and neutrinos. The muons are absorbed in steel and earth and the beam as it reaches the surface is solely of neutrinos which pass through the Big European Bubble Chamber (BEBC), two counter experi-

ments and the heavy-liquid bubble chamber Gargamelle.

- The protons can be directed to a second underground target to produce 75-GeV kaons or 110-GeV antiprotons which run into BEBC.

- The beam can be brought to the surface and split into three on entering the West Hall. Three targets are available to feed the twelve separate experiments set up in the West Hall with secondary beams of up to 150 GeV. The dimensions of the building—which existed prior to the SPS—preclude using higher energies; the North Area will be custom built for 400 GeV.

Why, it may be asked, has CERN built a 400-GeV proton accelerator when the Fermilab at Batavia, Illinois, installed a comparable machine more than four years ago—is it not needless duplication? The answer is complex and depends as much on subjective opinion as objective fact. But the list of SPS experiments approved at CERN gives some idea of the differences. Most striking are the number of neutrino experiments—13 out of 28. These take advantage of both the extremely high event rate anticipated from the neutrino beam-line and the excellent bubble chamber facilities. A second unique feature of CERN is the hyperon beam-line, yielding Ξ and Σ particles; an experiment on the decay characteristics of these particles will be watched with interest. A third will be the muon facility in the North Area.

It is also a mark of CERN's distinctive character that although roughly half of all experiments planned on the SPS are devoted to the "new physics"—hunts for J/ψ , intermediate vector bosons (maybe!), heavy leptons and charm—the other half tend to be "spectroscopic" experiments in the realms of "normal physics", painstakingly filling in gaps in knowledge without spectacular expectations.

Those who have worked in both Fermilab and CERN comment on the difference in style—Fermilab has a lot of flair, is able to latch on very quickly to new ideas, is vulnerable to failures (the magnets break down with

What SPS costs . . .

Item	Swiss Francs ¹ (millions)
Total CERN expenditure in 1975	647.9
of which SPS construction	237.9
SPS contract expenditure to October 1975: on	
site buildings and equipment	214.7
normal machinery components	194.6
special machinery components	89.7
Total	499.0
of which British contractors received	75.8
Projected total cost of building SPS (1970 prices)	1150

¹SFr 2.4 = \$1 SFr 4.0 = £1

. . . and who pays

Country	% in 1976	% in 1971
Austria	2.22	1.96
Belgium	4.02	3.77
Denmark	2.29	2.26
France	21.49	19.90
West Germany	25.40	23.27
Italy	13.34	12.89
Netherlands	5.30	4.43
Norway	1.60	1.52
Sweden	4.55	4.59
Switzerland	3.38	3.20
UK	16.41	21.61

Contributions are based on GNP

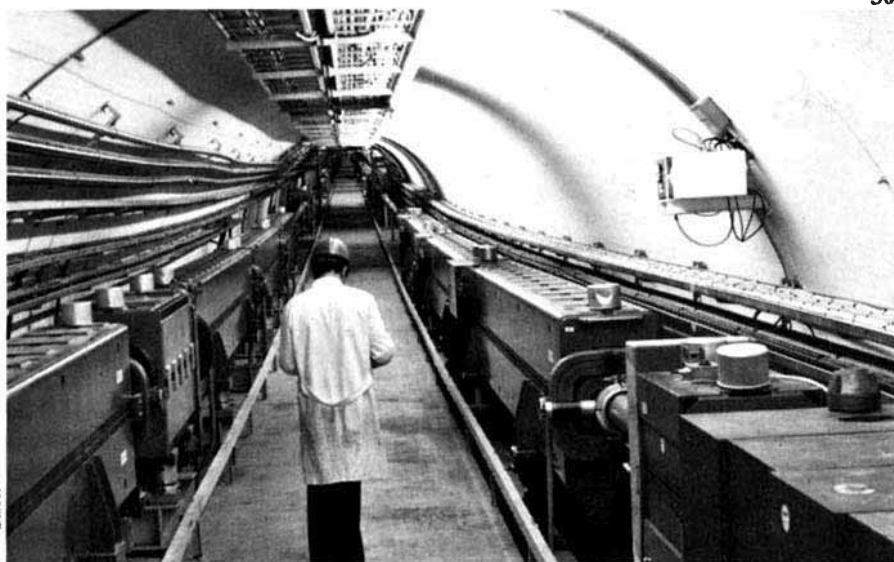
depressing regularity); CERN is more bureaucratic in its planning of experiments, is less prone to serendipity and yet runs so consistently well (SPS is expected to be extremely reliable) that much excellent definitive work can be done. As one physicist put it "at Fermilab you practically bring your experiments in cardboard boxes and stick them under tarpaulins; here look at the huge echoing halls and the gold-plated equipment!".

One of the problems that CERN is facing is the growing reliance of the high-energy physics community on one central facility as national facilities are tapered off. An exception is in Germany, where many reckon PETRA will produce some very exciting physics, and again will undoubtedly attract an international clientele; but otherwise there is nowhere else to go. Maybe the average number working on an experiment, which is 6 on the SC, 12 on the PS, 18 on the ISR and 25 on the SPS, reflects in part an increased complexity of experiment. But some believe that there are too many working on the SPS and that the quality of experience both for graduate students and research worker suffers accordingly. Another concern, raised first by Professor Jentschke last year when he was director of CERN's Laboratory I, is that very few physicists from CERN's smaller member states are finding their way into the teams for the SPS.

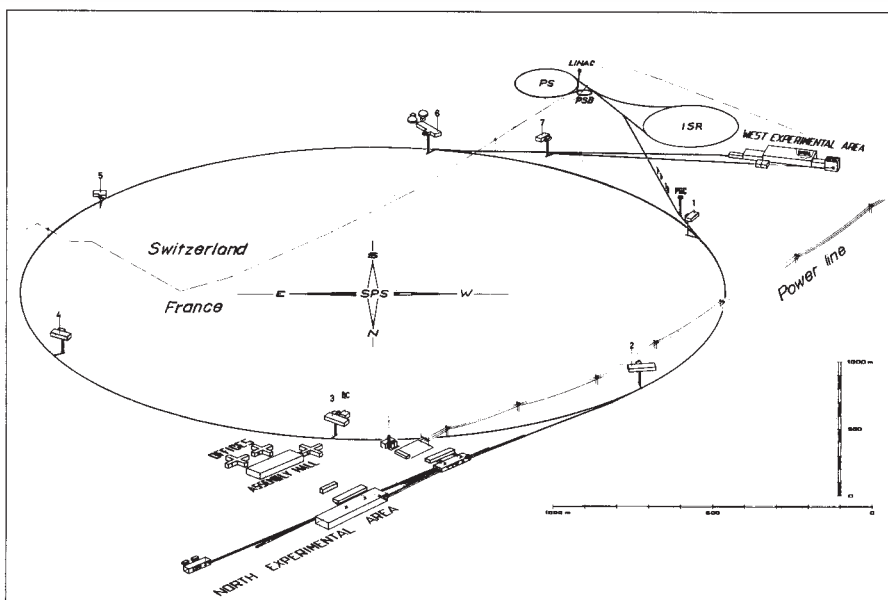
There is, on the other hand, some feeling around that CERN staff get a better deal than those who simply visit. The tax-free salaries are generous, to say the least, even allowing for the high cost of living in Geneva. And the cost of CERN's on-site services means that many (British, in particular) bring in their own technical staff.

Not surprisingly, there is a little apprehension at the moment at what financial adjustments the United Kingdom may try and negotiate at CERN's December Council. Dr J. B. Adams, the Director-General, is phlegmatic as always—"last year the Germans had some difficulties, the year before it was the French; these things come and go". But the real question is not a short term one. It is whether over the next five or ten years the nations contributing not only to CERN but also to other major accelerators are going to sustain their support for high-energy physics.

Within CERN itself there are already ideas circulating for new machines. The tunnelling for the SPS cost a relatively small fraction of the total bill, and some would like to see the SPS ultimately acquire an intersecting ring for proton/proton or even proton/antiproton studies. Such discussions,



Looking towards West Area from SPS



SPS plan, showing link with PS and experimental areas

however, are often coloured by a feeling that since the weak and electromagnetic interactions seem to be more fruitful sources of new states than strong interactions, a CERN venture into an electron machine might be a good next step. The European Committee for Future Accelerators has had a working party look at the physics potential of a 100-100 GeV electron-positron storage ring.

Fermilab, meanwhile, is exploring the possibilities of superconducting magnets to push its peak energy up to 1,000 GeV. The intellectual drive to get up to this level is that the postulated carrier of the weak force, the intermediate boson, probably has a mass so high (at least 37 GeV) that 1,000 GeV protons would be needed to unearth it. The prize from this could be unification of theories of weak and electromagnetic forces. CERN initially planned superconducting magnets in the SPS but eventually dropped back

to conventional ones because the technology was not far enough advanced.

At the same time Soviet physicists are looking to a 2,000 GeV proton synchrotron at the Institute of High Energy Physics, Serpukhov, as the next step from their 76-GeV synchrotron. Again, superconducting magnets are envisaged, to deliver the fields of up to 5 Teslas. Plans also include the simultaneous construction of a 20-GeV electron synchrotron with the capacity for proton-electron collisions.

It seems that plans for the next generation of machines are too far advanced in the United States and the Soviet Union for there to be any possibility of global collaboration. Attention has therefore turned to the idea of the next accelerator-but-one being a truly world machine, if only because a multi-billion-dollar price label is bound to be attached to it. A proton synchrotron at more than 10,000 GeV is a possible. But this is up to the politicians. □

USA

Where lies the test?

The Food and Drug Administration last week issued a set of regulations designed to help combat a growing scandal over drug testing in the United States. Colin Norman reports from Washington

A RELATIVELY unremarkable scientific paper, published early in 1972, has touched off a major scandal over the testing and regulation of drugs in the United States. It has led to allegations that some drug companies have conducted sloppy animal tests, faked results and presented misleading toxicity reports to the federal government—allegations which call into question the integrity of key parts of the drug regulation process and cast doubt on the safety of some products.

The allegations are so serious that the Food and Drug Administration (FDA) has already investigated the testing procedures and test reports of a few companies, and it is about to launch an unprecedented inquiry into the work of dozens of others. On the basis of what it has turned up so far, FDA last week issued a bulky set of regulations which essentially instruct the drug industry how to conduct animal tests and warn that drugs could be pulled off the market if it is subsequently discovered that misleading reports on them have been filed with FDA. The objective of the proposed new regulations, an FDA official said last week, is "to make sure that drug companies compete with each other on the quality of their professional work rather than on their level of cost-cutting".

The concern over the quality of drug testing was triggered by a study published in the *Journal of the National Cancer Institute* by Mario Rustia and Philippe Shubik of the Eppley Research Institute. The study indicated that a commonly prescribed drug, known as Flagyl, increases the incidence of pulmonary and lymphoid tumours in mice. The paper raised eyebrows in FDA because the agency had in its files reports of two studies submitted by the drug's manufacturer, G. D. Searle and Co., which claimed to show that Flagyl is not carcinogenic in rats.

A thorough check of Searle's study was then conducted by Adrian Gross, an FDA scientist. Gross concluded that Searle's study did indeed show evidence of carcinogenicity, but the company's report didn't reflect all the data. FDA officials took the matter up

with Searle, and two years later the company re-submitted its report. This time, Gross said in an interview last week, the summary agreed with the data, which would have been fine except that it was the data, not the summary, which had been altered. That finding prompted FDA to conduct a full-scale investigation into Searle's testing and reporting on several products. It was probably the most thorough federal investigation ever conducted of a drug company's scientific activities and it turned up some very disquieting findings.

The official report of the FDA investigation is among the most outspokenly critical federal documents that one could read. It states:

We have uncovered serious deficiencies in Searle's operations and practices which undermine the basis for reliance on Searle's integrity in conducting high quality animal research to accurately determine or characterise the toxic potential of its products . . . Searle has not submitted all the facts of experiments to FDA, retaining unto itself the unpermitted option of filtering, interpreting, and not submitting information which we would consider material to the safety evaluation of the product. Some of our findings suggest an attitude of disregard for FDA's mission of protection of the public health by selectively reporting the results of studies in a manner which allays the concerns of and questions of an FDA reviewer.

The task force which conducted the investigation recommended that the matter be turned over to the Justice Department, and that a grand jury be called to look into possible violation of the law. An FDA spokesman said last week that no grand jury investigation has yet been started, however. It should be noted that Searle's chief executive officer, Daniel Searle, acknowledged in Congressional testimony earlier this year that some "human errors" had occurred but he maintained that the FDA inquiry had found no evidence of fraud or misrepresentation by Searle personnel.

Whatever the outcome of the investigation, the Searle episode raises some very disturbing questions. FDA regulation of drugs and food additives is based largely on safety data submitted by the manufacturer of proposed new products. Thus, if manufacturers' test reports are unreliable, for any reason, much of the basis for FDA's regulatory actions could be open to doubt. And animal tests, it should be noted, are particularly important since they provide the chief means of identifying whether or not new products may have potential long-

term problems, such as carcinogenicity, teratogenicity, and mutagenicity—problems which don't show up in clinical trials on humans.

Because of the serious implications if FDA is regulating drugs and food additives on the basis of unreliable data, Congress earlier this year added \$16.4 million to the agency's budget for an investigation to determine whether the problems uncovered at Searle are widespread in the industry. Senator Edward Kennedy, who was instrumental in securing the extra funds and in bringing the matter to public attention, noted last January that "if data are proven false and misleading, then the regulatory decisions may be tragically wrong".

FDA's investigation is about to get under way in earnest. During the next three months, inspectors from the agency will look into testing procedures and facilities of about 40 drug companies and independent testing laboratories. But FDA has already gathered alarming evidence that the problems are not limited to Searle.

Last week, the agency proposed a set of so-called Good Laboratory Practice (GLP) regulations which spell out standards which FDA will require for animal tests of drugs and food additives. In a preamble to the proposed regulations, FDA Commissioner Alexander M. Schmidt states that a number of recent investigations conducted by FDA in cases where the agency had reason to be suspicious of test reports "have identified significant problems in the manner in which nonclinical laboratory studies are being performed". The deficiencies, Schmidt said, were uncovered during inspections of "major pharmaceutical firms, inspections of several private contract testing facilities, and internal reviews of toxicity studies of color additives conducted by FDA".

The problems included the following:

- "Pathology reports submitted to the agency were inconsistent with the original autopsy records".
- "Microscopic tissue slides were conducted by more than one pathologist, each of whom came to different conclusions, yet only the conclusions favourable to the drug were submitted to the agency".
- No records were kept for any individual animal in one long-term feeding study, "despite the fact that the records were required for proper analysis of the study and were represented to the agency to exist".
- "Animals were recorded as normal for a variety of factors, including

awareness, appearance, appetite, and thirst, when in fact the animals were dead".

- "Drugs were administered to animals in a manner which made it impossible to determine how much, if any, of the required dosage was actually ingested by the animal".

- In one toxicity study, "gross changes of tissue began to appear, yet management was not made aware of these alarming changes for approximately 4 to 8 months.

- In a study done by an independent testing laboratory, FDA was told that animal tissue samples had been examined when a review of laboratory records indicates that the tissue samples had never even been collected.

Such alarming, and possible fraudulent,

incidents are supposed to be stopped, or at least made more difficult, by FDA's proposed GLP regulations. But the regulations would in fact not be unduly burdensome. They set standards for the training of laboratory personnel, the quality of facilities and equipment, the handling of test animals, and the keeping and reporting of test data. They also specify that a Quality Assurance Unit must be established in each facility to ensure that the GLP regulations are followed. In short, the proposed standards are no more strict than those which should already be in force in any self-respecting facility.

Perhaps the most important aspect of the proposed regulations is that they spell out a number of sanctions which FDA could invoke if violations of the GLP standards are encountered. The ultimate sanction would be to

remove a product from the market if FDA had approved it on the basis of test reports subsequently found to be false or misleading. Another proposed sanction would be to disqualify a test facility as a source of information in support of a future marketing permit. Such action would be tantamount to shutting down a contract laboratory which makes its living from conducting toxicity tests for the drug industry. Finally, FDA points out that deliberate faking or misleading reporting of test results could be subject to criminal prosecution.

The proposed GLP regulations are open for comment for 120 days and FDA plans to conduct public hearings on them early next year. By that time, the results of the inspection effort should be available to provide an indication of the scope of the problem. □

SEVESO

The problems deepen

Italian authorities are now pondering how to clean up an area near Seveso in northern Italy, following an accident last July which released large quantities of dioxin into the environment. Some important new information bearing on the problem has surfaced in the United States. Colin Norman reports

DIOXIN in the soil of the most heavily contaminated area near Seveso can be expected to take up to 14 years to break down to levels at which the compound can no longer be detected. But there is unlikely to be much, if any, movement of dioxin in the soil, and the contamination will probably be confined to the upper 12 inches for a long time.

Those conclusions have emerged from an analysis of previously undisclosed studies conducted by the US Air Force, and from information gathered as a result of a bizarre dioxin contamination incident in Missouri in the early 1970s. The studies were ferreted out and analysed by Barry Commoner and Robert E. Scott, researchers at the Center for the Biology of Natural Systems at Washington University. Scott and Commoner have sent their conclusions to authorities investigating the Seveso disaster.

The Air Force studies, which were released to Commoner on October 28, involved soil tests with military herbicides contaminated with dioxin, left over from the Vietnam war. They involved heavy spraying of plots of land in Utah, Kansas and Florida, and subsequent monitoring of dioxin levels in

the soil. Commoner and Scott conclude from the studies that the likely half-life for breakdown of dioxin in the soil at Seveso is about 250 days, a rate which suggests that it would take about 14 years for contamination to disappear from the most heavily affected area (Zone A) and about seven years to degrade to undetectable levels (about 10 parts per trillion) in the surrounding zone (Zone B).

But, in spite of an annual rainfall of about 60 inches in Florida, the Air Force study indicated that dioxin in the test plot there did not migrate in the soil. That finding is particularly important because there has been recent heavy rainfall in Seveso which has led to worries that the dioxin may have been washed through the soil into the groundwater. "It may be expected that the TCDD present in the upper layers of the soil at Seveso will probably not migrate downward significantly, in spite of the heavy rainfall", Commoner and Scott conclude. Those observations lead them to note that if the soil at Seveso is left in place, contamination is unlikely to spread, but alternatively if the upper layers of soil are stripped off and replaced with fresh soil, any dioxin left in the lower layers would not be expected to migrate to the surface.

The Air Force study has also yielded information on the effects of low levels of dioxin on wildlife. Samples of wildlife around the Florida test plot indicated significant bioaccumulation of the contaminant, and there was also a marked drop in the fertility of beech mice in the area. The observations lead

Commoner and Scott to conclude that wildlife at Seveso will be affected unless soil concentrations can be reduced to about 500 parts per trillion, compared with the average concentration of about 10 parts per million in Zone A. Such a reduction, they note, "could be most rapidly achieved by removing the upper layer of soil and replacing it with uncontaminated soil".

The other information analysed by Commoner and Scott was derived from observations following an incident in which waste oil heavily contaminated with dioxin was sprayed on the ground in four places in Missouri in May and June, 1971. The dioxin contamination was unknown at the time.

One application involved spraying the ground in a horse arena shortly before a show to keep down dust. Hundreds of birds and several horses died as a result of exposure to dioxin and three months later a young girl who had been playing in the arena was admitted to hospital with symptoms ranging from nosebleeds and headaches to severe disuria and hematuria. In October, four months after the spraying, between six and eight inches of topsoil was removed from the area and replaced with fresh soil, but deaths of horses, cats and other wildlife persisted. In April, 1974, four years after the accident, a further six inches of soil was removed, whereupon the toxic effects disappeared.

The original contamination in the arena was reckoned to be about 30 parts per million, a level similar to that in Zone A, and thus the observations compiled by Commoner and Scott are directly relevant to the Seveso investigation.

Commoner and Scott conducted

their study under the auspices of the Scientists Institute for Public Information.

● **Alastair Hay adds:** At the time of the Seveso explosion, 150 women resident in the area were in the first trimester of pregnancy. Most of the women applied for a therapeutic abortion when informed of the teratogenic properties of dioxin. Following examination by health officials of 730 pregnant women formerly resident in the contaminated area, only 29 have had the operation. The fetuses removed from these women are currently being examined for any signs of abnormality.

Professor Francesco Cefis of the *Istituto Clinico di Perfezionamento* in Milan, and Professor Alfred Gropp, Head of the Pathology Department of Lübeck Medical School in Germany, are in charge of the investigations. Neither would comment last week on the results of their observations so far. Only five of the 29 fetuses have been examined and Professor Cefis estimated that it would take another two months to complete the study. By then women who were in the earlier stages of pregnancy in July will be six or seven months pregnant—too late for an abortion to be performed.

The rate of spontaneous abortions in the Seveso area is reported to have increased to twice the Italian national average since July. The rate of miscarriage was reported to have increased similarly in Vietnam following spraying with the herbicide 2,4,5-T. Both the trichlorophenol discharged at Seveso and the 2,4,5-T used in Vietnam contained dioxin. The Vietnamese authorities maintain that dioxin was responsible for the miscarriages

Towards anticipating disaster

In Britain last week the Insurance Technical Bureau (ITB), which leading insurance groups established in 1972 as a non-profit making organisation with "technical expertise in the field of industrial loss prevention", demonstrated 'Anticipator', a new system designed to alert managers of chemical plants to potential hazards in their site installations.

One of the recommendations made by the 1974 Court of Inquiry into the Flixborough disaster was that some form of recorder should be made available to monitor operating conditions in industrial installations. This would perform a function similar to that of the 'black box' flight recorder in aircraft. 'Anticipator' is a sophisticated monitoring system with its own computer which will register information about pressure, temperature, flow rates,

gas leaks, vibrations and other factors relevant to the safe operation of a chemical plant. It can be pre-programmed to record only data which fall outside normal working parameters, making interpretation a relatively simple process.

The designers argue that incidents usually "cast their shadows before them", and feel certain that Anticipator would have prevented accidents such as those at Flixborough and Seveso. With the primary research and development work now completed at a total cost of £50,000, the ITB propose to find a commercial concern to develop the system for marketing. The trouble is, the biggest firms do their own monitoring, and the smaller ones may not be able to afford the new device.

Alastair Hay

which they recorded.

Meanwhile, the scientific commission responsible for formulating a policy to decontaminate the area containing dioxin has still not agreed upon methods to be employed. Frustration at the apparent inactivity has been expressed both by experts anxious to begin the work of decontamination, and by former residents of the area who last week tore down protective fences in protest at the delay.

The work of the commission has been hampered by the disagreement among experts concerning the amount of dioxin to be dealt with. Dr Donald Lee of the Plant Pathology Laboratory at Harpenden, UK, suggests that 130 kg of dioxin was released in the explosion, and Professor Samuel Ep-

stein of the University of Illinois, considers the figure of 300 lb of dioxin to be "about right". But a group of chemical engineers invited by the commission to submit a programme for decontamination estimate that 1-3 kg of dioxin was released, a figure based on results obtained from a simulation of the explosion.

The two top officials of ICMESA, who after being arrested and charged with culpably causing a disaster were released on bail to supervise decontamination work within the ICMESA plant, have since been rearrested and criminal proceedings have begun. The criminal prosecution will be conducted at the same time as the civil proceedings brought against Hoffman-La Roche, ICMESA's owners, by the Italian authorities. □

EEC

Manoeuvring for agreement

The EEC's Council of Research Ministers met in Brussels last week to discuss once again the Joint European Torus (JET), the Community's fusion project, and the research programme for the Community's Joint Research Centre (JRC). Chris Sherwell reports

So near and yet so far. The EEC now looks sufficiently close to agreement on the site for JET that the small step the Council of Research Ministers took at their meeting last week may in retrospect come to look like a giant leap for the Community. But because there is still ample scope for more than one hitch, the prospect could be as much for a leap backwards as a leap for-

wards.

On paper the progress that everyone is keen to say really did occur in Brussels certainly looks very slight. The Council reached just one firm decision with any consequence in terms of action or expenditure. This was that the Community's fusion programme, which is undertaken by a number of associated laboratories in the member states but does not include JET, could go ahead for another four years.

To call that insignificant, though, is to reckon without the intricacies and nuances of Community deliberations. Until now there has been no agreement on expenditure for the full five years of the fusion programme. When the programme first came up for discussion

last December, Italy refused consent, saying she could not agree without a decision on a site for JET. She modified this position at a February Council meeting, allowing just one year's expenditure. In the view of one senior official in Brussels, therefore, the absence last week of continued further blocking of the programme from Italy means that the decision can be interpreted as a positive sign for JET.

What changed Italy's mind? The answer is to be found in the way the JET argument has developed. Italy's concern has always been principally over the fate of Ispra, one of the four establishments constituting the JRC—the others are in Germany, Belgium and the Netherlands, but Ispra is the largest with two-thirds of the JRC staff. By the last Council meeting in October it had become clear that progress on

JET might be possible if the Council would sanction in full the four-year (1977–1980) JRC research programme put up by the European Commission during the summer, much of which would go to Ispra.

The difficulty, however, was that Britain, France and Germany had made various demands for cuts in staff and expenditure on the proposed JRC programme. Germany was seeking cuts of less than 100 in staff. France was looking for cuts of around 200, and Britain had proposed a figure in excess of 300. Modifications in these demands was plainly necessary if there was to be any relaxation in the Italian demand for JET at Ispra—a demand which the Commission had itself backed on the grounds that a Community project should go to a Community centre.

The flexibility was evidently there, and the key development last week was a provisional agreement on these matters. Britain and France went to Brussels apparently prepared to pare back their demands to around 100—the number which it was thought could be lost through “natural wastage”—and eventually conceded along with Germany that the figure could be 80. This was the compromise proposal put up by the Commission after no agreement could be gained on the Commission's original compromise proposal of 50. It is damaging to Italy's hopes for Ispra inasmuch as the staff cuts, details of which obviously have to be finalised, are likely to fall most heavily on Ispra.

The agreement makes for a reduction in operating expenditure from 175 million units of account (mua) to 146 mua and so cuts the overall JRC budget for the four years to something like 346 mua. But there is a catch in all this: the agreement is conditional on

a suitable agreement being reached over JET. Thus, although the JRC research programme is at last final, little if anything can be done about it without further progress on the matter of the JET site—and on this there was little progress in Brussels.

What chinks of light there are, though, suggest that the matter is now delicately poised. Seven of the nine countries agreed last week that the site ought to be at a location with previous practical fusion and plasma physics experience. Significantly, it was France and Italy who thought this not to be relevant. France's proposed site, Cadarache, does not have fusion experience. Ispra has fusion experience and is not officially withdrawn, even though it now looks a less strong candidate with the provisional agreement on the JRC programme. The cases of both Culham in Britain and Garching in West Germany are strengthened by their experience.

The uncertainty is thus tangible. There are differences between France and the other countries on the degree of Euratom staffing that JET ought to have. But these are thought not to be insuperable. More importantly, the Council could not even agree last week on a procedure by which agreement might be reached. As a result, the Dutch chairman, Mr Brinkhorst, and the EEC Research Commissioner, Mr Brunner, are to tour the countries concerned to discuss the sites still further. If they determine that the basis for an agreement exists—the agreement when it comes will have to be unanimous—they will call for another Council meeting, provisionally scheduled for December 20.

The Community's fusion director, Mr Palumbo, has a few illusions about

that meeting, since if it takes place it would be the last one as far as JET was concerned. If it doesn't take place, he stresses, the whole project could be seriously endangered. Officials concerned with the JRC are somewhat more sanguine about the impact of a delay on the research programme, even though it is due to start in January. But the delay could be enormous. A new country takes the Council chair in January. That is Britain. And a new president takes over at the European Commission. He is from Britain. The latter change could well exacerbate any delay. Both changes could, in the present circumstances, make a difference to the outcome over JET.

No one, of course, is saying that the changes will help Culham win JET. But it was generally agreed last week that the choice had now narrowed substantially. In fact the case for Culham may well be strengthening. If Ispra is now discounted, the argument goes, that leaves Cadarache and Garching as Culham's competitors. And the rather sudden strength in the previously quietening voice for Cadarache can be discounted because it comes too late, because seven countries think experience is important and because Britain wouldn't wear it.

Garching can also be discounted, the argument goes on, because the Germans once again did not push strongly for it at last week's meeting, because it has the biggest fusion programme of the nine countries and plans its own large machine (AZTEC) anyway, and because it already has a Community research establishment. But if Germany does not get JET, what would it get? Well, speculate the cynics, watch out for a German filling the post of JET project director. □

BRITAIN

Changing the framework

A UK parliamentary committee last week published a 96-page report on university-industry relations. Chris Sherwell reports on the various institutional proposals of what could become a controversial document

WITH friends like the House of Commons Select Committee on Science and Technology, those amongst Britain's science community looking for a bit of stability in the country's science policy framework may not need enemies. For at the very time when the projected evolution of that framework seems virtually complete, the committee has published a wide-ranging report going beyond the narrowest confines of its

subject to offer proposals for further substantial changes.

The report is the result of written submissions and a score of public hearings initiated by the committee's Science Sub-committee. It follows an interim report on scientific research in British universities published in July 1975 and a second report released at the end of last year, and is “concerned with the purposes of the institutions of advanced scientific education and research”. The committee, clearly inspired by the idea that science policy should relate to the general social and economic objectives of the community, says it is essential “that we should be prepared to re-examine the organisation of science and scientific education in

terms of our current needs”.

Its recommendations regarding the framework within which British science policy is conducted may not throw the whole matter into the melting pot again. But in focusing directly on the country's capacity to conduct research and development they seem certain to fuel the arguments over whether an explanation of Britain's economic performance is to be found in the areas of innovation, investment and productivity, and hence also in the very platform of its science policy.

The most far-reaching recommendations are for the appointment of a junior minister in the Department of Education and Science (DES) with responsibility for science, for the transfer of funds from the Science Research Council (SRC) to the Department of Industry (DoI), for a review of the

relationship of basic and applied research in Britain, and for a possible new institution to replace the National Research Development Corporation (NRDC).

The proposal for a minister emerges from a discussion of education and training of engineers and applied scientists, for which the report argues most strongly. Among other things the committee also says the concept of Special Institutions for Scientific Education and Research (SISTERS) should be "revived and implemented", and urges the DoI and SRC to devote greater attention to the "teaching company" idea—a recommendation which post-dated by a couple of days an announcement of a £500,000 scheme for three such companies to improve university-industry links. But the idea of a Science Minister will look curious to many, and not just because it is an old idea that was dismissed when the present policy framework was worked out.

The minister, the report says, "should be principally concerned with scientific and technical education at all levels of the education system, and with the activities funded from the Science Budget". That seems unlikely to find much favour among those who relate the need for a minister to the amount of money he would handle, and still less likely if the biggest money-eater in the Science Budget, the SRC, ought according to the committee to have some of its funds transferred elsewhere.

The argument for that transfer—perhaps the most contentious proposal—is straightforward. The direction of a country's basic research, it goes, affects its capacity to do applied research which is geared to national needs. Traditional academic criteria directing basic research cannot guarantee that capacity, even with unlimited money; with declining funds, a research council system which uses such criteria is a luxury. Knowing this, the argument goes on, the research councils talk in terms of national needs and are put in an intolerable position; the SRC, whose support for applied research (unlike other research councils) is not subject to the ameliorating effects of the customer-contractor principle, suffers especially.

Thus, proclaims the report, "there is a good case for the transfer of a proportion of the Council's funds to the Department of Industry, which is the natural 'customer' department for the applied research supported by the SRC". The idea is to let the SRC take the decisions it is qualified to take, and not review research needs in terms of inadequately-defined national priorities. Those priorities should be determined by those responsible—the commissioning departments acting through their

ministers in Cabinet.

The proposal for a review of the relationship of basic and applied science is similarly related to the existing framework of science policy. In the three years up to April 1976 during which funds previously disposed of by research councils were transferred to the relevant departments under the customer-contractor arrangements, other organisational changes went ahead. Departments acquired chief scientists, and the emerging task of coordination between departments fell to an inter-departmental committee of chief scientists and permanent secretaries. The position of Chief Scientific Adviser to the government fell away, but the Cabinet's Central Policy Review Staff in turn had a chief scientist attached to it.

In addition another entirely new body was also created. Known as the Advisory Council for Applied Research and Development (ACARD), it was designed as the equivalent in applied research to the Advisory Board for the Research Councils (ABRC) in basic research; it is chaired by the Lord Privy Seal. Welcoming ACARD, the committee says it should "review the relationship between government-supported applied R&D and government-funded basic research as a matter of urgency". In particular it should "examine the operation of the customer-contractor relationship and of the ABRC to ensure that effective machinery exists for relating basic science policies to long term departmental R&D strategies". ACARD's reviews should be published, and the Lord Privy Seal ought also to make annual reports to parliament, the committee says. It does not consider whether ACARD is the appropriate body to conduct such examinations.

The committee claims it is not making sweeping recommendations for changes in the organisation of government R&D. Among its other proposals is one that the government undertake "a thorough review of the level and nature of the research undertaken in their own establishments" and attempt to transfer to universities and polytechnics "work of a more basic nature, not requiring major physical research facilities, wherever this is possible". Another is that encouragement be given, for example, to bringing higher education and industry into closer alignment; the committee says there is a good case for devising financial incentives "possibly in the form of generous tax allowances" to encourage companies to place research contracts with universities.

The most controversial points it makes in this area, though, are likely to be those concerning the NRDC. The committee says "urgent attention"

should be paid by the government to the "mis-match between the activities of the SRC and NRDC", and "urgent action" should be taken to correct it. It recommends a number of changes which, if implemented, would make the NRDC less concerned with producing a financial return from the results of research and more able to give advice and assistance. The functions it proposes for NRDC, says the Committee, may well be better performed "by a new institution without the accumulated scepticism and indifference which NRDC's policy and activities appear to have generated in some quarters"; as it is, the NRDC activities are "in no way conducive to encouraging the exploitation of academic research".

Last week the head of NRDC, comparatively unexcited by the report, said he would have been surprised if there hadn't been some adverse criticism, particularly as the committee provided an ideal forum for complaint. He happily acknowledged that there was some truth in the report's comments on NRDC but thought many of the criticisms were unjustified, adding that NRDC expenditure in relation to higher education institutions would have to be perhaps quadrupled if the right service was to be given, and more than a simple change in the NRDC's terms of reference was involved.

The SRC had not by the beginning of this week made any plans to issue a statement on the committee's report. Whatever the reaction to it inside universities and industry, it seems clear that its real impact will depend ultimately on how the Select Committee itself is viewed as an institution. The committee's members are hoping now for a parliamentary debate on the subject, which seems the only viable means by which its recommendations can be urged upon the government short of some positive response from outside. □

● A team of about 10 scientists would be adequate for British biological warfare research, a junior Defence Minister, Mr Gilbert, said in a parliamentary written reply last week. A reduction in military research in the Microbiological Research Establishment at Porton Down, Salisbury, had been on the cards since the announcement of a review of its spending in the March 1976 White Paper on Defence. Mr Gilbert's announcement means that the establishment's future depends on its civil work, at present under study by the government's Central Policy Review Staff and the National Institute for Biological Standards and Control, who hope to report before the end of the year. Up to one third of the establishment's running costs are currently met by revenue from its civil work.

news and views

Intermittent nerve conduction

from Shin-Ho Chung

ALMOST every axon in the central nervous system ramifies profusely before making synaptic contacts with its target cells. This terminal arborisation is formed by the axon bifurcating into branches, the branches into twigs, and the twigs into fine terminal filaments. A single axon entering the superior cervical ganglion, for example, establishes about 1,500 synapses with its post-synaptic neurones, and all of these contacts are believed to be activated by an impulse in the main axon. However, several recent studies have shown that not every impulse flowing down the main axon invades all of its terminals; instead, at each bifurcation some impulses are channelled into one branch and are prevented from invading the other. This process, known as 'intermittent nerve conduction', is incompatible with the classical idea of the axonal tree as a simple distributor of nerve impulses to a large number of the next-order neurones, and it invites further investigation.

Barron and Matthews (*J. Physiol. Lond.*, **85**, 75; 1935) first noticed the phenomenon of intermittent conduction in a class of sensory fibres in the cat, which enters the spinal cord and sends branches back to the skin through neighbouring dorsal roots. When monitored in a branch coming out of the cord, a steady stream of impulses entering the cord was found to be transformed into an intermittent series, with an alternating sequence of discharge and silence. The pattern of this conduction block depended on the frequency of discharge in the recording fibre as well as on the amount of neuronal activity in the neighbouring fibres; for example, when the spinal cord was active with ascending sensory discharges, the block became frequent and long-lasting. Barron and Matthews believed that this phenomenon of intermittent conduction was one of the mechanisms by which the central nervous system regulates the flow of impulse traffic and 'transforms

a simple sensory discharge into a sensation.' After the dorsal root reflex was described (Toennies, *J. Neurophysiol.*, **1**, 378; 1938), Barron and Matthews' observations were erroneously assumed to be a special instance of it, and, except for a few sporadic notes (Fuortes, *J. Physiol. Lond.*, **112**, 42p; 1950; Howland *et al.*, *J. Neurophysiol.*, **18**, 1; 1955), their very accurate and detailed descriptions were, in the main, ignored.

The phenomenon has now been re-examined in the simple nervous system of invertebrates, where the branching pattern of a single axon can be viewed under a microscope. In the crustacean leg or abdomen, for example, a single motor neurone branches many times in its intramuscular course, and the terminals of the two main branches innervate two separate groups of muscles (Bittner, *J. gen. Physiol.*, **51**, 731; 1968). When the main axon is stimulated at a low, steady frequency, impulses travel to both groups of muscles, but the junctional potential is large enough to cause contraction in only one of them. As the frequency of stimulation increases, the speed of propagation into the branch serving the contracting muscle becomes reduced as the impulse momentarily hesitates before invading it (Hatt and Smith, *J. Physiol. Lond.*, **259**, 367; 1976). The flow of impulses into this branch is then abruptly blocked, whereas the other branch continues to admit impulses and the muscle served by it begins to contract (Grossman *et al.*, *Brain Res.*, **64**, 379; 1973; Parnas, *J. Neurophysiol.*, **35**, 903; 1972). Hints of a similar phenomenon exist in the results of studies on a variety of other neurones such as the sensory afferent in the leech (Van Essen, *J. Physiol. Lond.*, **230**, 509; 1973) and the giant axon of the cockroach (Spira *et al.*, *J. Neurophysiol.*, **39**, 882; 1976; Castel *et al.*, *ibid.*, 900).

The mechanism by which the propagation of nerve impulses is blocked

is not thoroughly understood. There are several reasons for supposing that the conduction failure occurs at the point of branching or at the region where the geometry of the axon undergoes a sudden change. Ordinarily, the currents produced by an impulse travelling down an unbranched axon have about four times the magnitude necessary to ensure continuous invasion. At points of branching, the current from the parent axon is split between the daughter branches and, at the same time, must charge a larger capacitance. This causes not only an attenuation of the exciting current, but also a hesitation of the nerve impulse, allowing sodium inactivation to develop. Such a region with a low safety margin is likely to be most susceptible to a small shift in ionic environment, and accumulation of some ions in the extracellular space following repetitive stimulation may further reduce the safety margin so that the transmission of nerve impulses fails completely (Smith and Hatt, *J. Neurophysiol.*, **39**, 794; 1976; Spira *et al.*, *op. cit.*). Theoretical computations based on the Hodgkin-Huxley equations predict that the increased concentration of potassium in the extracellular space during repetitive firing will cause conduction failure at the region where the geometry of the axon undergoes a sudden constriction (Parnas *et al.*, *J. Neurophysiol.*, **39**, 909; 1976). This, however, is not the only condition which may bring about the conduction block. Prolonged hyperpolarisation of the axon (Van Essen, *op. cit.*), temporary anoxia (Krnjevic and Miledi, *J. Physiol. Lond.*, **149**, 1; 1959), or currents arising from activity of neighbouring fibres (Lettvin *et al.*, *Brain Behav. Evol.*, **3**, 72; 1970) may have the same effect.

Whatever the precise mechanism ultimately turns out to be, the finding that axonal trees do not faithfully convey impulses to their terminals will have several important implications in

neurobiology. If we accept the classical view of the axon as a passive distributor of impulses, then there appears to be a serious defect in its performance. In order to overcome this defect, the nervous system may form a large number of redundant connections so that intermittent failure of nerve conduction will not seriously hinder the processing of neural information. Alternatively, as Lettvin and his colleagues maintain (*Brain Behav. Evol.*, *op. cit.*: see also Waxman, *Brain Res.*, **47**, 269; 1972), the

axonal tree may act as a filter, regulating the flow of impulses into one or other branch, depending on the temporal pattern of firing. It is quite common for a single nerve to innervate more than one type of post-synaptic cell, and if the axon is capable of controlling one or the other, or both, through its firing pattern, as demonstrated in the crustacean motorneuron, then what appears to be a fatal defect of the bifurcating axon may turn out to be a very useful property in achieving an economy in neural organisation.

Mysteries of early dinosaur evolution

from Barry Cox

ALTHOUGH many pages have been written discussing the mystery of the extinction of the dinosaurs, almost as much uncertainty surrounds their origin—or origins. The large Mesozoic archosaurian reptiles known as dinosaurs belong to two separate groups, the Saurischia and the Ornithischia, which differ in such characters as the structure of the pelvic girdle and the possession of a predentary bone (present only in ornithischians). A diversity of saurischians is known from the late Triassic, as are about six genera of ornithischians. But the relationships of the two groups to each other, or to the earlier Triassic archosaurians known as pseudosuchians, remain uncertain.

By the Early Jurassic, saurischians include at least two types. The small carnivores (or coelurosaurs) and the large carnivores (or carnosaurs) were both bipedal, and are both placed in a single group, the Theropoda. Late Triassic coelurosaurs are known, but no Triassic carnosaurs have yet been found. The other saurischian group, the Sauropoda, comprises the very large, quadrupedal, long-necked herbivores of the Jurassic and Cretaceous. Their Triassic ancestry runs back into an assemblage of quite large facultatively bipedal forms, the prosauropods, which include both herbivorous and carnivorous types. It is therefore at least possible that the coelurosaurs, the carnosaurs and the prosauropods all had quite separate origins from within the pseudosuchians. Various authors, such as Charig and Bonaparte, have suggested this, basing their views mainly on apparently divergent structural modifications of the ankle in these groups.

So far, at least, there has been less evidence of diversity in the Triassic ornithischians; only two, rather similar families are known. Smaller than most of the saurischians, their structure has

been less well known, but an article by Santa Luca, Crompton and Charig in this week's issue of *Nature* (page 324) provides a preliminary account of an exceptionally fine specimen from the very late Triassic of South Africa. This little bipedal, fast-running herbivore, *Heterodontosaurus*, had already developed a number of specialised features, such as the grasping hand and some degree of fusion of the bones of the ankle. Ornithischians, then, had also clearly been evolving and radiating for some time before the end of the Triassic.

If one now attempts to relate these two types of dinosaur to the Triassic pseudosuchians, there appears to be a puzzling overlap in time between the two groups, although possible evolutionary links between them obstinately refuse to appear. Though the pseudosuchians include large bipedal carnivores, the ornithosuchids of the Middle and Late Triassic, which have been suggested as ancestral to the carnosaurian dinosaurs (Walker, *Phil. Trans. R. Soc.*, **B248**, 53; 1964), this has since been denied by Bonaparte (*Coll. int. Cent. nat. Rech. sci.*, **218**, 485; 1975). The large Late Triassic quadrupedal pseudosuchian carnivores known as rauisuchids (or prestosuchids) have been nominated as the possible source of the saurischians, while the only pseudosuchian similar to the little ornithischian dinosaurs is the small bipedal genus *Euparkeria* of the Early Triassic.

The possible ancestors of the different types of dinosaurs are thus scattered through the Triassic, and the pseudosuchians persisted until the end of that Period. Throughout the Late Triassic, at least, a variety of pseudosuchians therefore coexisted with a variety of their presumed descendants, the dinosaurs, as has been pointed out by Sill (*Bull. Mus. comp. Zoo. Harv.*, **146**, 317; 1974). The perhaps multiple

evolution of dinosaur lineages therefore did not lead to the rapid ecological replacement of their pseudosuchian ancestors. These instead persisted, together with their relatives the semi-aquatic phytosaurs and the pig-like rooting aetosaurs, until the end of the Triassic, when all became extinct together—though no new ecological replacements for them appeared at that time.

So the poor palaeontologist searching for answers is therefore, in the origin of the dinosaurs, confronted with complexity where he hoped for simplicity, while in the replacement of the pseudosuchians by their varied offspring he meets a sudden (if delayed) simple event where he expected complexity! Today, the fashionable *deus ex machina* for explaining extinctions is continental drift, but even rising sea levels resulting from the beginning of the opening of the Atlantic in the Late Triassic seem unlikely to provide a solution to these puzzles. □

Future of the CERN intersecting storage rings

from M. G. Albrow

A Workshop on Future ISR Physics was held at CERN, Geneva on October 4–15, 1976.

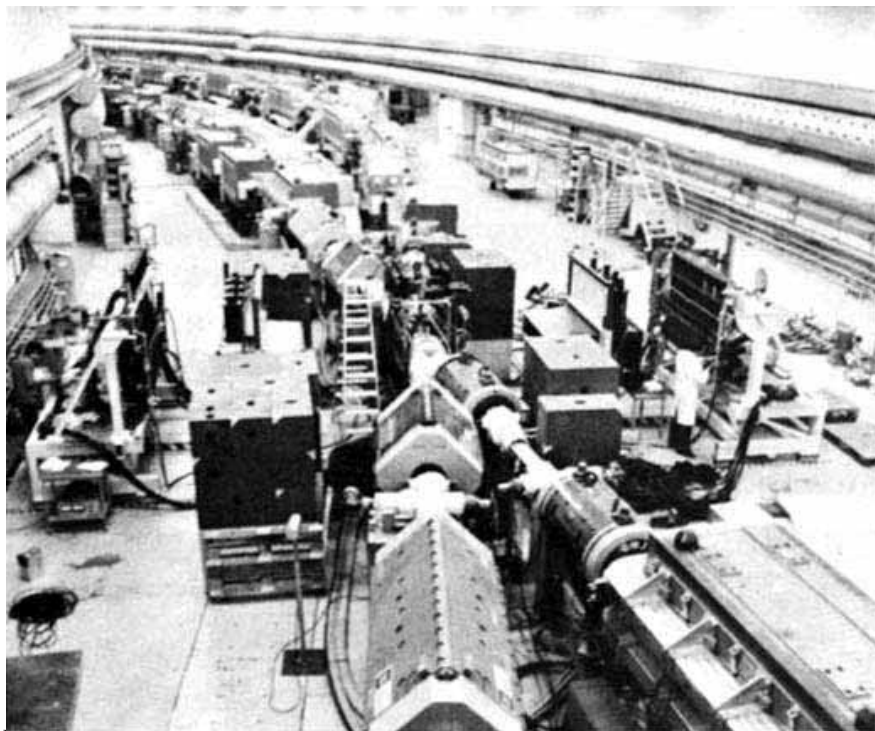
IN January 1971 two beams of high energy protons were trapped in the CERN Intersecting Storage Rings (ISR) and brought into head-on collision for the first time. This achievement enabled high energy physicists to study particle interactions, under well-controlled laboratory conditions, at energies otherwise accessible only with cosmic rays. The ISR beam energy ranges up to 31 GeV: an accelerator of 2,000 GeV would be needed to obtain the same effective collision energy using a particle beam on a stationary target. The present commissioning of the 400 GeV SPS accelerator at CERN, and plans at the Fermi National Accelerator Laboratory to construct a 1,000 GeV "Energy Doubler", have stimulated a re-evaluation of the role of the ISR in high energy physics. There is no doubt that the ISR has a unique position in terms of available collision energy; the question is whether that advantage is important enough to guarantee an exciting programme of physics for many years to come, and therefore to justify its continued operation in the present financial climate. A workshop, attended

by about 50 physicists from Europe and the United States, was recently held at CERN to discuss such questions and try to define a research programme for the ISR after the present generation of experiments, that is, into the next decade. Here I have attempted to summarise the most promising of the ideas considered.

The concept that the strongly-interacting particles (or hadrons) are composite objects containing more fundamental constituents, for example quarks and antiquarks, has achieved almost universal recognition over the past 12 years. All the known hadrons can be constructed from a basic set of four quarks q (up, down, strange and charmed) and their four antiquarks $\bar{q}$. The proton is supposed to contain three "valence" quarks (up, up and down) which determine its quantum numbers, as well as a sea of $q\bar{q}$ pairs and particles called gluons which bind the quarks together. In this picture, when two protons pass through each other the point-like quarks usually miss each other—they may however get shaken up with excitation energy. These excited protons then break up into normal ground state hadrons. Occasionally a $q\bar{q}$ or a $q\bar{q}$ collision occurs, and because of their point-like nature a large angle scattering is likely. The scattered quarks themselves do not emerge as free particles, but "dress up" into normal hadrons which frequently have large momentum components transverse to the collision axis. The discovery at the ISR of relatively large numbers of hadrons with large transverse momenta supports this concept of point-like hadron constituents, just as the observation of large angle α particle atom scattering led Rutherford to discover the atomic nucleus.

However much more data are needed on this rather rare type of collision to establish the precise nature of quark quark scattering, and the next generation of experiments probably requires a large new particle detection system, which would allow the identification and momentum measurement of most of the particles created in the collisions. The importance of the ISR in this field is that these processes have a very strong energy dependence and are therefore less amenable to study in lower energy machines.

Many of the present experiments at the ISR concentrate on lepton production. Four leptons are known (e^- , μ^- , ν_e , ν_μ and their antiparticles). It may be more than a coincidence that we know of four quarks. The leptons do not interact strongly and should only be produced indirectly in hadron-hadron collisions, for example through the production and decay of



One of the intersection regions of the ISR showing the point (centre) at which the two regions cross.

other particles (such as charmed hadrons, see *News and Views*, **262**, 537; 1976) or through $q\bar{q}$ annihilation:

$$q\bar{q} \rightarrow \gamma \rightarrow \mu^+ \mu^-$$

where γ is a virtual photon or a heavy photon-like particle, for instance the J/ψ . (Samuel Ting and Burton Richter were awarded the 1976 Nobel prize for physics for the discovery of this particle, see *Nature*, **263**, 714; 1976.) Current experiments are searching for new J/ψ -like particles and, if the annihilation mechanism is correct, are able to map out the q and $\bar{q}$ distributions inside the proton. Many physicists expect new surprises in this field; if so the research programme is completely open ended.

The high energy (and wide energy range) of the ISR is also vital for the study of some other, perhaps less fashionable, processes. I mentioned earlier the possibility of one proton becoming excited into a higher energy state in a collision; the other proton may recoil in its ground state acting only as a "spectator". The energy or mass of these excited states grows linearly with the effective collision energy, and reaches values of order 12 GeV at the ISR (compared with about 3–4 GeV at the SPS and Fermilab). Little is known about these massive states, for example their composition in terms of particle type. A related process is expected to "turn on" in the ISR energy range, in which both incident protons emerge un-

changed from the collision except for a small loss of energy. The energy lost by the protons reappears in the form of hadrons which essentially "pop out" of the vacuum (the vacuum is supposed to be full of virtual particle pairs which can materialise if sufficient energy is supplied). If this process does not exist several theoretical ideas must be incorrect; if it does exist it is rather fundamental and should be studied in depth.

The ISR Workshop also discussed possible improvements to the machine itself: higher collision rates, other particles in the beams (antiprotons, deuterons and heavier ions) and higher energy beams. Beams of 120 GeV would be possible if the present magnets were replaced with 5 Tesla superconducting magnets, giving an energy equivalent to that of a 30,000 GeV accelerator. This is clearly a possibility for longer term development, and in my view a most exciting one. Theories which unify the electromagnetic and weak forces predict three very fundamental particles ($W^\pm$, Z^0) which are responsible for the weak interaction (radioactivity) with masses near 65 and 80 GeV. They would probably be produced at the ISR if it were upgraded in this way, and much of the other physics studied at present would greatly benefit from an extension in energy.

I believe most of the participants left the Workshop with considerable enthusiasm for the future possibilities of the ISR. First, there are many questions about very high energy inter-

actions that need answering, and which appear to be answerable with the machine as it is today, but are beyond the range of lower energy machines. Second, improvements to the machine and the detectors will widen the scope of these investigations, and it is not unreasonable to suppose that new surprising discoveries will be made. Finally the investment made in the ISR has given Europe a unique facility which could be made a stepping-stone into a completely new energy region. □

Pair-breaking in superfluid ^3He

from P. V. E. McClintock

STUDIES of the way in which negative ions move through liquid ^3He have resulted in a particularly revealing demonstration of the superfluid properties which it exhibits in the temperature range around 1 mK. The experiments, carried out at Helsinki Technical University by A. I. Ahonen, J. Kokko, O. V. Lounasmaa, M. A. Paalanen, R. C. Richardson, W. Schoepe and Y. Takano, and reported in *Physical Review Letters* (37, 511; 1976), represent the first successful investigation of ion motion in the recently discovered superfluid phases of the liquid.

The so-called ion formed when an excess electron is injected into liquid helium is actually a semi-macroscopic spherical structure with a diameter of around 3 nm and an effective mass about a hundred times greater than that of a helium atom. Being negatively charged, ions may be directed by means of externally applied electric fields and their arrival at a collecting electrode may be detected as a current. They therefore constitute particularly convenient probes for elucidating the behaviour of the liquid, and they have already been employed to excellent effect in gaining insights into the nature and fundamental properties of superfluid ^4He . Comparable investigations of superfluid ^3He are, of course, considerably more difficult because of the exceedingly low temperatures which must be achieved and maintained: the superfluid transition temperature of ^3He is near 2 mK, or about a thousand times colder than that of ^4He .

The ^3He in the Helsinki experiment was refrigerated, using a nuclear demagnetisation cryostat, to temperatures less than half that of the superfluid transition. The measurements themselves, based on a technique recently developed at Lancaster University, were in essence exceedingly simple. Bunches of ions, injected into the liquid from a field emission tip,

were gated into a drift space of known length which they traversed under the influence of a fairly uniform electric field, finally arriving at a collecting electrode where they were detected as pulses of current. By measuring the time taken to cross the drift space, the characteristic velocity of the ions could thus be determined for a number of different temperatures and electric fields.

Some typical experimental results for the B phase (the form of superfluid ^3He which is stable at low pressure) are shown in Fig. 1. It is strikingly evident that, for any given electric field, the ions travel much faster as the liquid is cooled below the temperature T_c of the superfluid transition. By analogy with the behaviour of ions in liquid ^4He , it is natural to try and account for the general shape of these curves on the assumption that there are two separate types of mechanism by which the kinetic energy of a moving ion may be dissipated: (1) at all finite temperatures the ion travels through the residual gas of normal, that is, unpaired, ^3He atoms which tend to impede its progress; (2) if the velocity of the ion is high enough, additional dissipative mechanisms can come into play whereby the ion creates excitations in the liquid. On this model, the ionic drift velocity which is measured in practice represents an average over the numerous individual scattering and possible excitation creation events which occur during the transit.

For sufficiently small ionic drift velocities, where only (1) need be considered, the velocity-field characteristic is expected to be linear with its gradient defining the ionic mobility. It may be noted from the figure that this is indeed the case and that, as the temperature is reduced, the mobility increases rapidly, consistent with the expected decrease in scattering as more and more of the ^3He atoms fall into the paired states that comprise the superfluid. Indeed, at the lowest temperatures, the mobility apparently became so large that it was impossible to observe the linear region at all. Comparing these results quantitatively with theoretical predictions by T. Soda (in *Proc. 14th Int. Conf. on Low Temp. Phys.*, edit by Krusius, M. and Vuorio, M., 1, 13, North-Holland, Amsterdam, 1975) and R. M. Bowley (*J. Phys.*, C9, L151; 1976), Ahonen *et al.* point out that both of these calculations substantially underestimate the magnitude of the superfluid mobilities. Probably, however, we will not have to wait very long for the theorists to improve their calculations, now that they have the stimulus of some real experimental data.

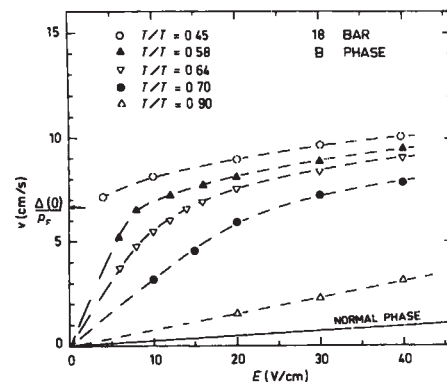


Fig. 1 The drift velocity v of negative ions moving through superfluid ^3He under the influence of an electric field E , for various temperatures. The behaviour in normal (non-superfluid) ^3He is shown by the full line, and the predicted value of Landau critical velocity for pair-breaking is indicated by an arrow.

Assuming that the observed bending over of the characteristics for velocities above a few cm s^{-1} is the result of mechanisms of type (2), it is natural to enquire into the nature of the excitations which are being created, and there are several possible candidates. For each of these there will be a so-called Landau critical velocity v_L , which the ion must exceed if the general requirement for simultaneous conservation of energy and momentum is to be satisfied during the creation process. At very low temperatures, where dissipation through (1) may be ignored, one would expect that an ion moving under the influence of a small electric field will accelerate without hindrance until it reaches v_L , and that it will then continue moving at a constant drift velocity equal to v_L while dissipating, in the form of excitations, the energy gained from the electric field. Thus, the position where the lowest temperature curve in the figure, that for $T=0.45 T_c$, cuts the velocity axis gives a value of v_L , which can be used to try and deduce the nature of the excitations. The simplest assumption, by analogy with the superfluid electron gas in a superconductor which in many respects is a similar system, is that the excitations emitted by the ion are just unpaired ^3He atoms. The Landau critical velocity for breaking apart the pairs of ^3He atoms forming the superfluid may be estimated as about $v_L=6.7 \text{ cm s}^{-1}$, and is indicated by an arrow in the figure. The fact that the curve looks as though it would extrapolate back to cut the velocity axis very close to this point is thus of considerable interest and significance, implying that energy dissipation under these conditions is indeed through the anticipated pair-breaking mechanism, rather than through crea-

tion of the (vortex-like?) excitations which may have given rise to the much lower critical velocities observed in the flow experiment recently performed by R. M. Mueller, E. B. Flint and E. D. Adams (*Phys. Rev. Lett.*, **36**, 1460; 1976) at the University of Florida.

Although a number of important conclusions may be drawn from the present data, the Helsinki experiment—as the authors themselves point out—is really only a beginning. Their principal achievement is to have demonstrated the technical feasibility of using ions to probe the superfluid phases of ^3He . It is to be hoped that future experiments of this type, taken in conjunction with improvements in the theoretical situation, will enable the temperature dependence of the energy gap separating the paired atoms of the superfluid from the (unpaired) excited states of the liquid to be mapped out in detail; it should be possible, using a more carefully designed geometry, to explore the anisotropy of the A phase; and the use of positive ions, to judge from earlier experience in superfluid ^4He , may perhaps enable completely new types of excitation to be created and identified. □

Nitrogen aggregates in diamond

from John Walker

As regular readers of News and Views may remember, nitrogen is a common impurity in diamond, though how exactly it got there and the precise forms it adopts within the crystal are only poorly understood. But a recent paper by G. Davies (*J. Phys. C: Solid St. Phys.*, **9**, L537–L542; 1976) partly resolves the latter question. The question is important, and not only for scientific reasons since it bears on the synthesis of diamond.

Synthetic diamonds are usually yellow, because they contain a small amount (perhaps one part per million) of dispersed nitrogen. That is, each nitrogen atom is isolated from the other nitrogen atoms, and replaces one carbon atom in the diamond lattice. In the current classification these are type Ib diamonds. Natural diamonds of this type are rare—one in a thousand.

The majority of natural diamonds, and all the large, well-formed ones suitable for adorning sceptres and film stars' fingers, are what we call type Ia stones. They also contain nitrogen impurity, in quite large proportions (up to almost one per cent); but in contrast to type Ib the nitrogen is aggregated. It appears that in growing her diamonds

Nature did not use the same process as General Electric and De Beers (fortunately for the gem industry, since yellow diamonds, though quite pretty, lack the sparkle and fire of water-white stones).

So what is the form of these aggregates? Until a couple of years ago we thought we knew the answer. The nitrogen gathered in large clusters, many thousands of angstroms in diameter, along {100} planes—the “platelets” that Evans and Phaal (*Proc. R. Soc. Lond.*, **270**, 538–552; 1962) had discovered by electron microscopy. Unfortunately we were wrong. We do not know what the platelets are, but thanks to earlier work by Davies (*Nature*, **228**, 758; 1970) and by E. V. Sobolev *et al.* (*Sov. Phys. Dokl.*, **12**, 665–668; 1967) we know that they contain little if any nitrogen.

We do know that there are at least two different types of aggregate, the so-called A and B forms, which give rise to characteristic infrared and ultraviolet absorption. What Davies has done in his most recent paper is to study one of the A bands, in fact a sharp line, under uniaxial stress. That is, he has squeezed a rectangular diamond block between two anvils, at stresses of the order of 20,000 atmospheres. As Wedlake has remarked, for diamond this is only a romantic squeeze; but it can nonetheless produce interesting results. Although the diamond lattice is cubic, under stress its symmetry is reduced (to tetragonal, trigonal or rhombic depending on the direction of stress), causing absorption bands to split into two or more components (just like the Zeeman effect in atomic spectroscopy—the splitting of spectral lines by a magnetic field).

Davies has deduced from the splitting pattern that the defect responsible for the A form of nitrogen has trigonal symmetry. The simplest atomic arrangement consistent with this is a pair of nitrogen atoms next to each other (“nearest neighbours”), replacing a pair of carbon atoms in the diamond lattice. Such a model fits in with the other constraints placed by current experimental knowledge: it is not paramagnetic, it is too small to be detected by electron microscopy, and it accounts for the slight increase in lattice parameter observed in nitrogen-bearing diamonds compared with pure specimens.

In a further paper (*Proc. R. Soc. Lond.*, **A351**, 245–265; 1976) Davies and his colleagues, again using uniaxial stress, demonstrate that the well-known H3 absorption and luminescence system in irradiated and annealed diamond has rhombic I symmetry. Davies (*J. Phys. C: Solid St. Phys.*, **5**, 2534–2542; 1972) has already shown that the H3 defect is produced when the A form of nitrogen traps a vacancy or interstitial

carbon atom produced during irradiation. Rhombic symmetry follows neatly from the trigonal symmetry of the A aggregate.

What we do not yet know is did the nitrogen aggregate into pairs by diffusion, or was it created like that during synthesis? What is the atomic arrangement of the B aggregate, and is nitrogen involved in the formation of the platelets? We still have many questions left to answer. □

Electron transfer systems in microorganisms

from D. O. Hall

An International Symposium on Electron Transfer Systems in Microorganisms was held at the CNRS Laboratory of Chemical Bacteriology, Marseilles on November 2–5, 1976.

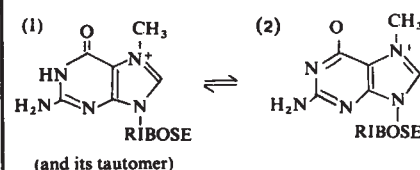
THE Marseilles laboratory has been active in the biochemistry of electron transfer in anaerobic bacteria for many years and the meeting was held to discuss the requirements of electron transfer at both the cellular and the molecular level.

The genetics of yeast metabolism has been a fruitful subject for research for some time. P. Slonimski (Laboratory of Molecular Genetics, Paris) discussed some of the recently discovered mitochondrial mutants (MIT[−]) which result in deficiencies in various aspects of mitochondrial metabolism. More than a thousand mutants are now available, opening up the possibility that mitochondrial functions for many different electron transfer reagents and enzymes will be available to biochemists to study the details of electron transport and ATP formation.

B. Haddock (University of Dundee) described his work on electron transfer systems in *E. coli* and *Paracoccus denitrificans*. It is becoming clear that

Correction

In the article “Eukaryotic mRNA: trouble at the 5'-end” (*Nature*, **263**, 188; 1976) the formula on page 190 was printed incorrectly. The correct formula is given below.



bacteria possess an electron transport chain containing many of the components familiar from mitochondria, such as flavoproteins, iron-sulphur proteins, ubiquinones and cytochromes. Since these bacteria are simpler to work with we now have, for example, a means of studying proton translocation to verify aspects of the Mitchell chemiosmotic theory and also to look at the role of compounds such as cyclic AMP in determining alternative pathways of electron transfer. Both these organisms can be grown in anaerobic and aerobic conditions, with precisely specified nutritional conditions, sulphate or iron limitation for example, which will affect components of the electron transport chain.

The occurrence of the enzyme superoxide dismutase to protect organisms against the harmful effect of the superoxide which is produced when oxygen interacts with free radicals is now a very active field. Only recently have anaerobic organisms been shown to contain superoxide dismutase. C. Hatchikian (CNRS, Marseilles) reported on the isolation and purification of superoxide dismutase from the sulphate-reducing organism *Desulphovibrio*, an obligate anaerobe. This enzyme has iron in its active centre and is similar to the superoxide dismutase from aerobic prokaryotes, such as *E. coli* and the blue-green alga *Spirulina*. One of the hypotheses put forward as to why anaerobes need such an enzyme is that so-called anaerobic organisms need to be tolerant to small quantities of oxygen. The superoxide dismutase may be inherent to these organisms, or may have been acquired from an aerobic organism by plasmid transfer.

I. Probst (Göttingen) reported on the interesting symbiotic relationship between the sulphate-reducing, anaerobic bacterium *Desulphuromonas* and a green, sulphur photosynthetic bacterium *Chlorobium*. *Desulphuromonas* can utilise acetate as its electron donor and sulphur as its electron acceptor; the net energy difference in this reaction is rather small, but is quite sufficient for the organism to grow in symbiotic association with green photosynthetic bacteria. The green photosynthetic organism *Chlorobium* uses CO_2 in a classic photosynthetic mechanism, with H_2S as its electron donor, and produces sulphur, which is then used by *Desulphuromonas* with acetate as its carbon source to produce CO_2 and H_2S thus completing the cycle.

The recent discovery of three different ferredoxins was reported by A. Xavier and J. Moura (Lisbon Institute of Science and Technology) from *Desulphovibrio gigas*. These ferredoxins exist as trimers or tetramers and interestingly enough are able to catalyse electron transfer at negative

and positive redox potentials in states of oxidation similar to those of the so-called "C-states" of Carter. What Xavier and Moura have shown is that naturally occurring 4Fe-4S proteins are able to exist as oligomers which themselves can give variations in redox potential. If it is a natural phenomenon and not simply induced by association of monomers *in vitro*, this may be the way in which iron-sulphur proteins can alter their redox potential either within a specific monomer or within the molecule as a whole by association of oligomers; oligomers, of course, can be induced by alterations in salt concentration, pH, and so on. This may in fact be one of the ways in which complex iron-sulphur proteins like nitrogenase and hydrogenase may be able to accumulate a number of pairs of charges before they catalyse reactions such as nitrogen fixation and hydrogen evolution. □

Control of external radiation dose

from a Correspondent

An international meeting sponsored by the Society for Radiological Protection and the UKAEA was held on October 19, 1976, at AERE Harwell.

THE protection of the environment has been widely discussed over the past few months. The objective of this meeting was to consider the main sources of occupational exposure which come from external β , γ and neutron radiations. J. A. Dennis (NRPB, Harwell) discussed the concept of dose equivalent index and related topics. He expressed the personal opinion that the present lack of guidance from the International Commission on Radiation Units on the use of surveying parameters could give rise to overexposures. It is necessary to measure the dose equivalent at different depths in the body and the NRPB has proposed depths of 5 to 10 mg cm^{-2} and 700 mg cm^{-2} for skin and body respectively. J. R. Harvey (CEGB, Berkeley) was keen to introduce the concept of dose equivalent ceiling to avoid any possibility of overexposure. Thus it would be possible that a survey would give a potential dose of 200 mrem in a particular situation whereas measurements with personal dosimeters would give 100 mrem, both being equally valid, and it is on the latter dose that ultimate control would be exercised.

F. W. Spiers (University of Leeds) described the work by himself and T. Ashton to measure the dose to different organs in a man phantom, exposed to isotropic γ -ray sources (energy range 0.05 to 2 MeV). During discussion Spiers suggested that the dose to the active bone marrow was a reasonable approximation to the whole body dose and is approximately the dose under 53 mm of tissue. He also stated that the response obtained by rotating a phantom in a beam produces different results from those obtained by the use of isotropic radiation.

The large differences in shielding required by nuclear reactors and high energy accelerators was brought out by two papers by A. F. Avery (AEA, Winfrith) and G. R. Stevenson (CERN II, Switzerland). The accelerator shielding consists of earth and masses of iron (100 m long by 2 m in diameter) to bring the dose rate down to 1 mrem h^{-1} for occupational locations and 85 $\mu\text{rem h}^{-1}$ at the site fence. Nuclear reactor shielding is required to reduce the dose rate to 0.1 mrem h^{-1} in these areas, but of course far less material is needed.

In contrast to these two papers E. J. Henshaw (Radiological Protection Centre, Liverpool) said radiation levels up to 1 rem h^{-1} were common in fluoroscopy, but these levels are only present for brief periods. Henshaw suggested that it was important to measure the X-ray generator parameters, in particular the potential across the tube. During the discussion it was suggested that it would be of interest to do a comparative cost-benefit analysis between reactor shielding and dose control in radiology departments.

H. W. Julius (TNO, Netherlands) contrasted the film dosimeter and the thermoluminescent dosimeter (TLD) for personal dose control. The former provides much more detailed information but the TLD has advantages of low atomic number, and linearity of response with fewer environmental effects. J. A. Douglas (AERE, Harwell) described some of the work of Cavallini from Italy and his own studies on the limitations of the albedo technique for personal neutron dosimetry. His work had showed that the ratio of the incident to reflected flux of neutrons could not be used to give the energy of the neutrons except in certain circumstances which could apply for a limited range of radiation shields. J. R. A. Lakey (Royal Naval College, London) brought the meeting to a close by discussing what we do when we cannot reduce the dose rate below 2.5 mrem h^{-1} . He emphasised the importance of working schedules and the sharing of the dose between workers. □

articles

Slip-line field theory and large-scale continental tectonics

Paul Tapponnier

Laboratoire de Geologie Structurale, U. S. T. L., Montpellier, France

Peter Molnar

Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A simple analogy is made between the tectonics of Asia and deformation in a rigidly indented rigid-plastic solid. India is analogous to the indenter and the great strike-slip faults correspond to slip lines. For various indentation geometries, the sense and linearity (or curvature) of strike-slip faults, convergence at the Burma arc and the existence of the Himalayan Burman Syntax, the conjugate strike-slip faults in Mongolia and the extension at the Baikal and Shansi graben can be predicted. Given the horizontal force necessary to support Tibet, an average shear stress of a few to several hundred bars along faults in Asia is predicted, corresponding to the yield stress of rigid-plastic material.

THE pattern of faulting in Asia, as revealed by analyses of Landsat photos and of seismic data (Fig. 1), bears a strong resemblance to the slip-line field calculated^{1,2} and observed³ for plane indentation of a plastic medium by a rigid die⁴⁻⁶ (Fig. 2). The analogous pattern of deformation for these two seemingly unrelated situations suggests that the continuum approach and the physical understanding of indentation of plastic materials, developed by mechanical engineers, could be used to gain further insight into the process by which continents deform when they collide with one another.

Rigid-plastic materials and plane strain slip-line fields

A plastic material is often assumed to be homogeneous, isotropic and perfectly plastic. There is no volume change and no strain hardening. A large class of problems, and those of interest here, assume a state of plane strain for which the two common criteria for the onset of plastic flow, or yielding (Tresca's and Von Mises's), are equivalent. Yielding occurs when the maximum shear stress, $\tau = (\sigma_1 - \sigma_3)/2$, reaches a limiting value, the yield stress, τ_y , where σ_1 and σ_3 are the maximum and minimum compressive stresses, respectively. Because of the mathematical difficulties in matching boundary conditions across the boundary between regions of elastic and plastic deformation and because the deformation in the plastic region is much greater than that in the elastic region, the elastic region is often treated as a rigid solid. Hence the material is said to be rigid-plastic: plastic where the yield stress is reached and rigid elsewhere. Figure 3a shows the stress-strain curve for such a solid.

All analyses are made at the yield point, so that $\tau \leq \tau_y$ throughout the region which deforms plastically. In this plastic

region the slip lines correspond to the maximum shear stress trajectories along which $\tau = \tau_y$. The mean (hydrostatic) stress, $\sigma_h = (\sigma_1 + \sigma_3)/2$, changes along the slip lines when they are curved, in a simple manner that can be determined from the equilibrium (Hencky) equations. Thus, from these equations, the geometry of the slip lines and the stress boundary conditions, the complete stress field in the plastic medium can be obtained easily.

Pure shear occurs on and parallel to the slip lines. They mark lines across which the tangential component of displacement can be discontinuous, as at faults in the Earth. The two types, α and β lines, differ only in sense, and in geologic terminology are right- and left-lateral respectively. As slip lines are lines of maximum shear stress, they must intersect each other at right angles and all traction-free boundaries at angles of 45° .

Solutions to rigid-plastic problems can be divided into those of steady and of unsteady flow¹. For steady flow, the boundaries do not change shape or relative position, and the pattern of flow is independent of time. The drawing of wire is an example. For most tectonic problems, however, it appears that solutions only for unsteady flow are of interest. Again two types of solution can occur. In some cases, the deformation occurs in such a way that, although the region of deformation increases with time and, hence, the boundary conditions change continuously, the geometry of deformation changes only in scale. An example is the indentation of a rigid-plastic medium by an infinite rigid wedge (Fig. 2b). The pattern of deformation changes only by increasing in extent with time, proportionally to the rate of indentation. For most problems, however, the actual shape of the boundaries, and, hence, the boundary conditions, change as deformation progresses. The plane indentation solution shown in Fig. 2a is an example; the slip lines are appropriate only for the onset of yielding.

In principle, it is possible to determine the finite deformation of a rigid-plastic medium by successively and incrementally calculating the deformation for the given boundary conditions and re-evaluating the boundary conditions on the deformed boundaries. For the analogy with Asian tectonics, however, we do not know with confidence the detailed history of deformation, and therefore sophisticated computation of complicated deformation histories probably will not be very useful at present. On the other hand, we think we do know, qualitatively, how deformation occurred during the past few million years, and we are more concerned with understanding it than speculating on the previous history of deformation. Fortunately, because the solutions for plasticity problems are incremental, in so far as strain hardening can be neglected, knowledge of the previous history of deformation is not necessary for calculation of the slip lines at a given stage of the deformation.

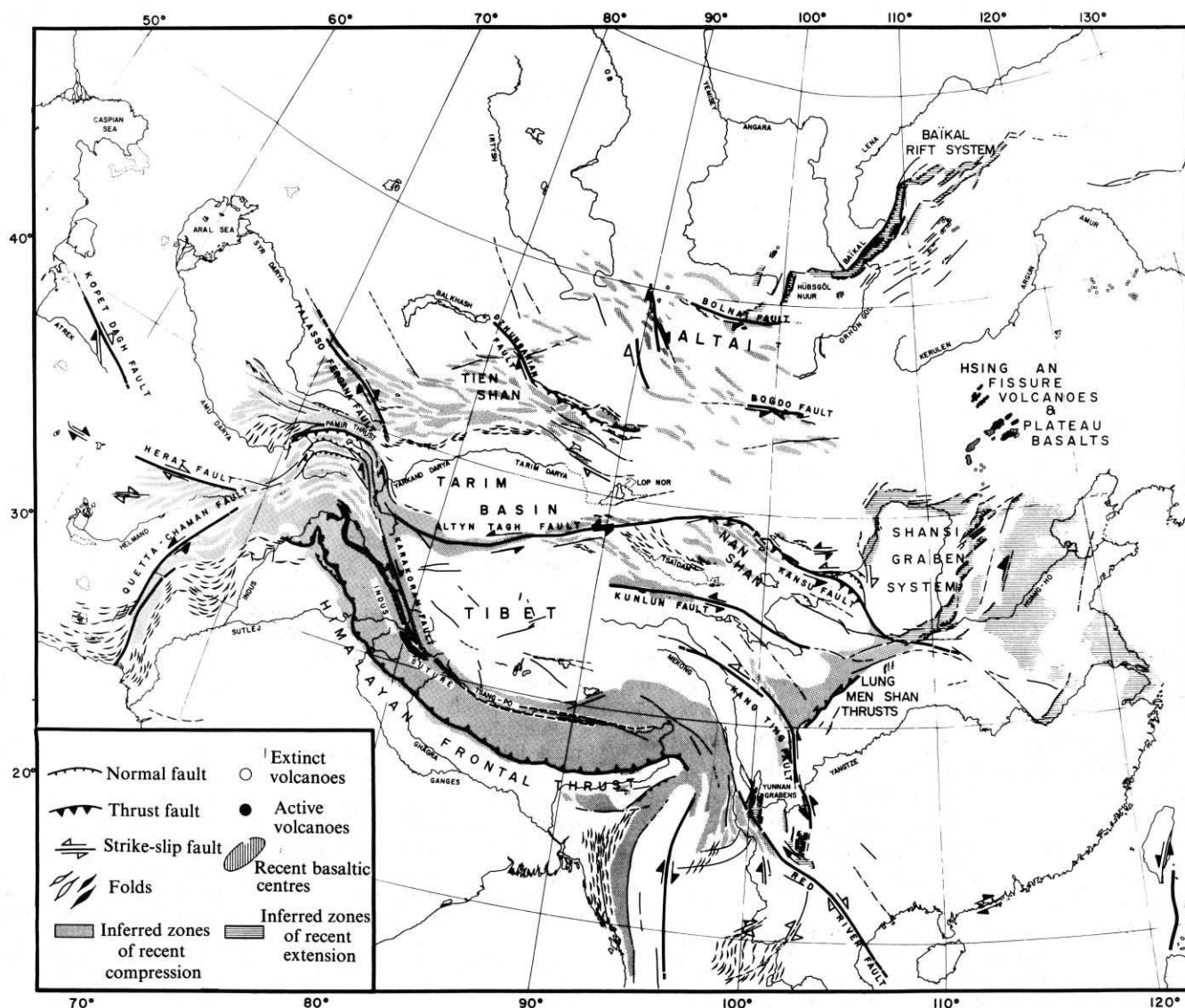


Fig. 1 Simplified map of recent tectonics in Asia⁵. (—) Major faults; (—) sense of motion; (▨) region of crustal thickening; (▧) regions of crustal thinning.

Assumption of plasticity and plane strain in Asian tectonics

Although in detail there are obvious violations of the basic assumptions made for the solutions for slip lines of rigidly indented rigid-plastic media when this formalism is applied to Asia, the similarity of the patterns alone is enough to pursue the analogy further. Although strain may accumulate elastically during intervals between earthquakes, the total strain in Asia since the collision of India is a few tens of per cent⁶. Elastic strain at any instant and the total change of volume in the region will be negligible portions of the total deformation.

If the stress on and near active faults alternately accumulates and then drops during earthquakes, the regional stress-strain curve may resemble that shown in Fig. 3b. Although the stress may vary with time, when averaged over a long time (many earthquakes), the regional stress-strain curve would be similar to that shown in Fig. 3a, for a perfectly rigid-plastic solid. In the Earth, at slow strain rates, strain hardening may be a minor effect, particularly along fault zones where gouge alteration and strain heating are likely to occur. Moreover, at depths ≥ 20 km fault creep is likely to play an increasingly important role in the deformation of the crust. At even greater depths, the temperature is appropriate for plastic deformation to be

dominant in most rocks at grain scale. Thus if most of the lithosphere at depth behaves plastically, in spite of the pressure dependent strength of the rocks in the top part of the crust, the assumptions of rigid plasticity and isovolumetric deformation are likely to be reasonable over a large scale and long time (10^6 yr) for the behaviour of the continental lithosphere as a whole.

Also, a plane horizontal strain analysis is justified for most of the deformation in Asia in the past 40×10^6 yr since the largest displacements (of the order of several hundred km (refs 5, 6)) seem to be horizontal movements along great vertical strike-slip faults.

Plane horizontal strain, however, does not take place everywhere in Asia. Important crustal thickening occurs at the Himalayas and Pamirs, and to a lesser degree in the Tien Shan and Nan Shan (Fig. 1). Further to the north-east (Fig. 1), crustal thinning occurs in the Baikal and Shansi graben belts. The deformation in these narrow mountain and graben belts, in fact, is close to plane vertical strain⁷⁻⁹, but is likely to be a fraction of that due to large scale strike-slip faulting, for two reasons: first, whereas strike-slip faulting is compatible with steady state deformation for long periods of time, in any given area crustal thinning is limited by the thickness of the crust itself and crustal thickening by the fact that, with increasing elevation

of the mountains, an increasing amount of work must be done against gravity. Second, the horizontal maximum principal stress necessary to cause plane strain ($\sigma_1 = \sigma_3 - 2\tau_y$) is less for plane horizontal strain through strike-slip faulting where $\sigma_3 < \sigma_2 - \sigma_1$ than that for plane vertical strain through thrusting and crustal thickening where $\sigma_3 = \sigma_2$. Crustal thickening will occur when σ_1 is especially high as is the case in front of the indenter or where there are pre-existing weak fault zones of the right orientation, as might be the case in old orogenic belts such as the Tien Shan or the Nan Shan (Fig. 1). In fact, that the crust and upper mantle structure of Asia is not homogeneous and may not be isotropic even on a scale of a few tens of km may be one of the causes of the differences between the fault pattern in Asia and the slip-line fields.

Indentation geometry and Asian tectonics

The deformation of a rigid-plastic body caused by indentation depends strongly on the shape of the indenter. If the indenter is flat, deformation extends into the semi-infinite rigid-plastic medium to a distance approximately equal to the width of the

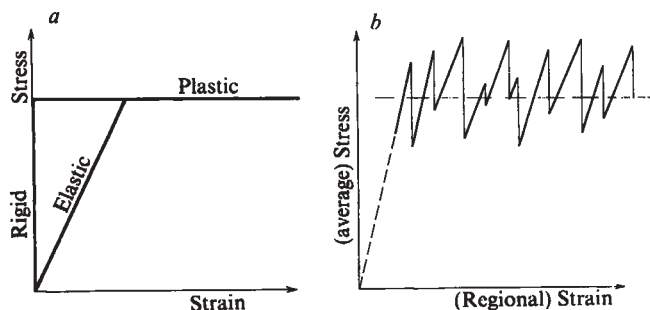


Fig. 3 *a*, Deviant stress-strain curve for rigid-plastic and elastic-plastic body. Once the yield stress, τ_y , is reached, plastic strain proceeds at constant stress (no strain hardening). *b*, Qualitative regional stress-strain behaviour of rocks in top part of crust. Slow accumulation of elastic strain is followed by sudden stress drops during earthquakes. With increasing regional strain, average deviant stress may oscillate about roughly constant value.

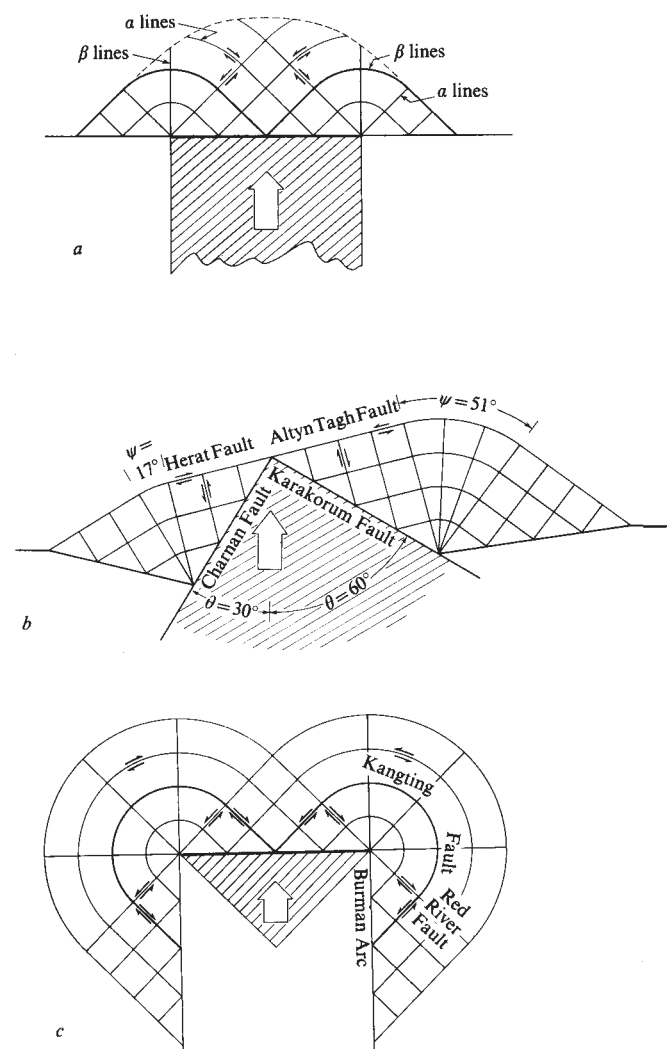


Fig. 2 Plane indentation of semi-infinite rigid-plastic media by different rigid dies. Arrows indicate sense of shear along slip lines. Principal stresses σ_1 and σ_3 bisect small quadrilaterals delineated by slip lines. The names of corresponding major tectonic features are indicated. *a*, Flat rigid die; *b*, rigid wedge—similar to the situation that occurs at the Pamirs (Fig. 1) (western end of the Himalayas); *c*, flat triangular indenter and hollowed-out medium, similar to situation which arises at the Himalaya-Burman syntaxis (Fig. 1) (eastern end of the Himalayas).

indenter (Fig. 2*a*). If the indenter is a wedge, at any given time the deformation may or may not extend further into the medium than the wedge itself depending upon the angle of the wedge (Fig. 2*b*).

Assuming the boundary of (rigid) India to be along the Himalayan Front and through the Kirthar and Sulaiman Ranges in Pakistan, parallel to the Chaman Fault, it seems that India has penetrated far into (plastic) Asia (Fig. 1), and therefore the analogy⁵ in Fig. 2*a* can be improved upon. In the western Himalayas and in Pakistan, the boundary of India is similar to that of a wedge. Note that the Herat and Altyn Tagh faults are approximately parallel to the slip lines that would be created by wedge indentation (Fig. 2*b*). As the direction of motion between India and Eurasia is approximately north-south, the western edge of the Indian wedge is steeper than the eastern edge. Correspondingly, the Altyn Tagh fault strikes more north-easterly than the Herat fault. The right lateral North Anatolian fault in Turkey⁴ and the left lateral Great Kabavi fault in Iran¹⁰ have an analogous relationship to Arabia, which appears to act as a wedge into Eurasia further west.

This simple analogy would predict strike-slip motion along the boundaries of the wedge and the rigid-plastic medium (Fig. 2*b*). Strike-slip motion of the right sense is observed for both the Chaman fault and the Dead Sea fault. There is also evidence for strike-slip faulting along the Zagros suture zone¹¹ and the Karakorum fault^{5,12}. In addition, however, thrust faulting and crustal thickening occur in both the Zagros and the Himalayas^{4,5,13}. The direction of thrust faulting is approximately parallel to the orientation of the maximum compressive stress in the plastic material (Fig. 2*b*).

An obvious feature of Asian tectonics is that deformation occurs over a vast area north-east of the Himalayas. If the wedge indentation analogy held for all of Asia, we might expect deformation north-east of the Himalayas to be bounded approximately by the Altyn Tagh fault, just as deformation does not extend far north-west of the Herat fault. A possible explanation for the widespread deformation north-east of the Himalayas lies in the fact that the trend of the Himalayas to the east curves and becomes approximately perpendicular to the direction of relative motion of the plates. Thus the analogy with the indentation solutions in Fig. 2*a* would be more applicable to this region.

We think that the pattern of faulting north, east and even south-east of the eastern Himalayas supports the analogy with plane indentation by a flat indenter. The Kunlun and Kangting faults are approximately parallel to β lines. The marked curvature, particularly of the Kangting fault, compared with the straight Altyn Tagh and Herat faults would be expected from their proximities to the edge of a flat indenter (Fig. 2*a*) and to a wedge (Fig. 2*b*) respectively.

For the region at the eastern end of the Himalayas an even

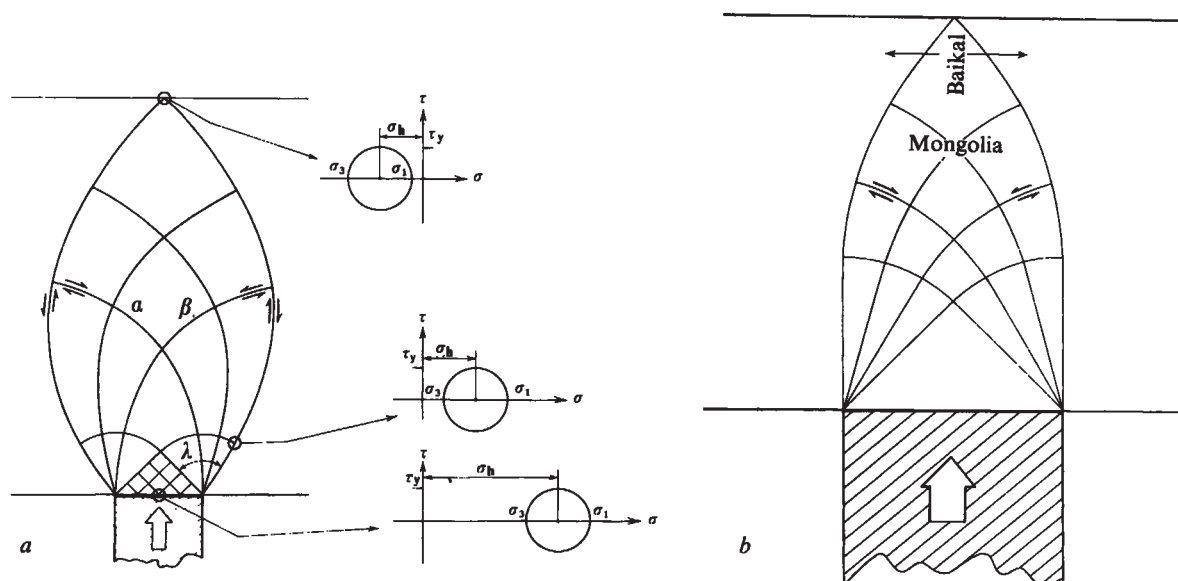


Fig. 4 Plane indentation of a bounded rigid-plastic medium². *a*, Width of plastic block is ~ 4.4 times width of indenter. Mohr circles on right show stress state at different points of field. σ_h decreases away from indenter as function of angle λ travelled along slip lines and becomes negative near boundary opposite indenter. *b*, Secondary tension may be cause of crustal extension at Baikal rift system.

more appropriate analogy than plane indentation of a flat-sided rigid-plastic medium is with the indentation of an already indented or a hollowed-out medium (Fig. 2c). For this case, the slip lines continue to curve around 180° in order to intersect the 'north-south' boundaries at 45° . In our interpretation of the Landsat photos, the Kangting fault curves around so that east of the Himalayas it trends north-south before becoming lost in the complications of the tectonics further to the south^{5,6}. In addition, the right lateral Red River fault^{5,6} is approximately parallel to an α line. Although this fault is longer than might be expected from the analogy in Fig. 2c, this analogy is strongly supported because it predicts not only the curvature and sense of motion on the Kangting fault but also the sense of motion and trend of the Red River fault.

For us a particularly puzzling feature was the Burman arc, in front of which are a series of folds and east-directed thrusts

and beneath which lies a belt of intermediate-depth earthquakes. Both observations imply a recent eastwards underthrusting beneath the arc. Thus, underthrusting of India beneath the Himalayas in a northerly direction and beneath the Burman arc to the east or south-east occurred simultaneously. In the indentation problem in Fig. 2c, material flows around the edge of the indenter and back in towards it. By analogy, we would expect India to squeeze part of Asia out of the way, so that this part of Asia extrudes around the edge of India and then encroaches upon its eastern side. As the Indian subcontinent is very narrow in eastern India between the Bay of Bengal and the Himalayas, oceanic lithosphere would have underthrust the Burman arc. Therefore resistance to the encroachment of extruded Asia would be less than if it encountered continental lithosphere. The Himalayan-Burman syntaxis and the surrounding tectonics are analogous to a corner in an indenter and the

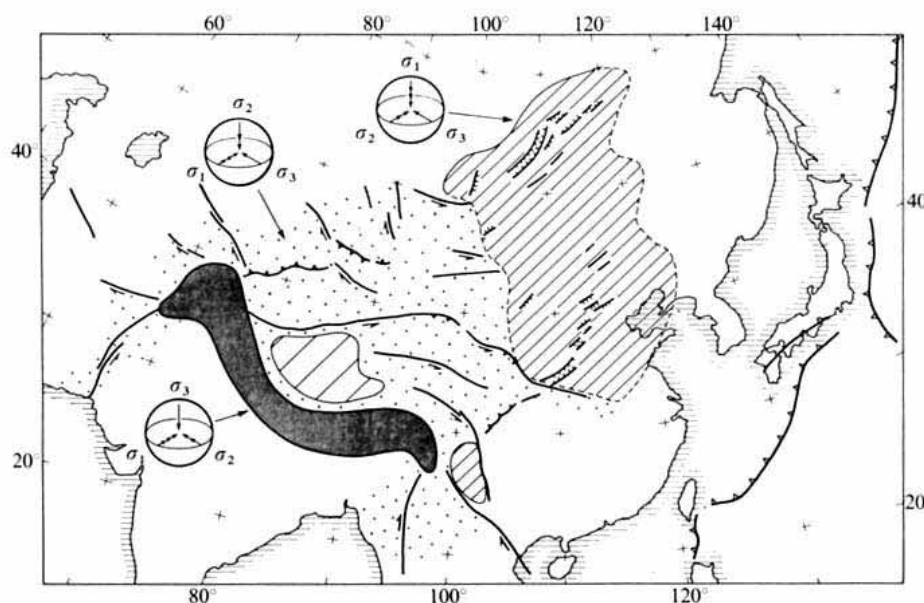


Fig. 5 Distribution of tectonic styles in Asia. Bold lines represent major faults (Fig. 1) (open thrust fault symbol for Pacific subduction zones). Dark shaded area—region of major crustal thickening, dotted area regions—where strike-slip faulting occurs, hatched area—normal faulting and crustal thinning. Corresponding stress states are indicated.

surrounding deformation when the indenter penetrates deeply into a rigid-plastic medium.

Boundary conditions at the limits of the indented block

The pattern of deformation in plane plastic strain also depends upon the shape of the boundaries, of the rigid-plastic medium, as is already clear from the differences in slip-line fields for the situations shown in Fig. 2a and c. For plane indentation as shown in Fig. 2, it is assumed that the rigid-plastic medium extends to very large distances compared with the width of the indenter. When a plastic block of finite width is indented, the slip-line field is of the type shown in Fig. 4a. Plastic yielding occurs in the region between the indenter and the symmetric pair of slip lines that leave the corners of the indenter and intersect each other at the opposite side of the block. Thus the plastic region becomes wider as the ratio of the width of the block (h) to the width of the indenter (a) increases (Fig. 4a). When this ratio reaches ~ 4.4 , the deformation changes drastically to resemble that in Fig. 2a; material flows around the edges of the indenter instead of forcing the two parts of the block to separate rigidly away from a narrow zone of plastic flow between them². Because the lithospheric plates are bounded, the slip-line field in Fig. 4a might be more appropriate for some problems than that in Fig. 2a. In a perfectly plastic material, the slip-line field will presumably be that either in Fig. 2a or in Fig. 4a, but not both, and not a combination of them. In the Earth, it is possible that elements of both situations might find analogues simultaneously, both because the continents are not homogeneous and because the shape of the plate boundaries and the boundary conditions are more complicated than in the idealised situations of Figs 2 and 4.

Perhaps the most interesting characteristic of slip-line fields of the type shown in Fig. 4a is that a tensile state of stress (secondary tension²) develops in the region adjacent to the boundary opposite the indenter. The cause of this tension is the

decrease of σ_h in the plastic medium away from the indenter. This decrease is expressed in Hencky's second theorem¹⁻³

$$\sigma_h = \sigma_{h(i)} - 2\tau\lambda$$

where $\sigma_{h(i)}$ is the maximum hydrostatic pressure in front of the indenter and λ the angle rotated clockwise along β lines and anticlockwise along α lines (Fig. 4). Along with σ_h both principal stresses σ_1 and σ_3 decrease proportionately to λ (Fig. 4a). In the extreme case shown in Fig. 4a, where $h/a \simeq 4.4$, σ_h is reduced enough near the apex of the field that all stress components are tensile. σ_3 becomes negative already in the middle of the field. This basic pattern is observed for all slip-line fields of the class shown in Fig. 4.

If the analogy of plane horizontal strain was strictly applicable to the Earth, then at a given depth the lithostatic stress σ_z would be equal to the intermediate stress $\sigma_2 = \sigma_h$. In the Earth, σ_z need not always be equal to σ_h , and when it is not, plane horizontal strain should not occur. There are two extreme situations in which this might occur: when $\sigma_z > \sigma_1 = \sigma_h + \tau_v$ or when $\sigma_z < \sigma_3 = \sigma_h - \tau_v$. These are likely to happen far from the indenter or close to it, respectively. Far from the indenter, the stress state will be appropriate for crustal thinning and the formation of graben (Figs. 4b and 5). We think that the Baikal Rift and Shansi grabens are a consequence of the stress system in Fig. 5, which is analogous to that in the area of secondary tension in Fig. 4. The predominantly normal faulting, the direction of extension and the relative position of the area in which it occurs to the Himalayas are appropriate to the conditions in Fig. 4b.

Deformation near the indenter, and mountain building

Near the indenter, σ_h is maximum and can be very large. When indentation starts it is likely that it will be large enough that σ_z will be the least compressive stress. Consequently, thrust faulting and crustal thickening will take place in the Earth. This in turn will lead to an increase in elevation and therefore to an increase in σ_z at a given depth below sea level. Plane horizontal strain will not take place until σ_z is large enough to be approximately the intermediate principal stress. At this stage, deformation may proceed through strike-slip faulting. This pattern of a first major phase of deformation with shallow thrusts and folds and a late phase of predominantly strike-slip faulting is the basic pattern to be found in most ancient orogenic belts⁸.

When plane horizontal strain is the dominant mode of deformation, then the mean elevation in any given area should reflect the local value of σ_h . The high altitude of Tibet and the general decrease in elevation away from Tibet in Asia may be a manifestation of the decrease in σ_h . Thus we think that both the distribution of tectonic styles in Asia (Fig. 5) and the decrease in mean elevation away from the Indian indenter may reflect a variation of σ_h similar to that predicted by the analogy in Fig. 4a.

The yield stress of Asia

For simple indentation geometries, the pressure σ_1 applied by the indenter and the yield stress are simply related. For the case in Fig. 2a, $\sigma_1 = 2\tau_v(1 + \pi/2)$; for that in Fig. 2, $\sigma_1 = 2\tau_v(1 + \pi)$; and for wedge indentation (Fig. 2b), $\sigma_1 = 2\tau_v(1 + \psi)$, where ψ is related to the wedge half angle θ by (see ref. 1)

$$\cos(2\theta - \psi) = \frac{\cos\psi}{1 + \sin\psi}$$

For $\theta = \pi/4$, $\psi \simeq 0.575$.

Assuming Tibet to be in isostatic equilibrium, we can estimate σ_1 for different depths of compensation. Because Tibet

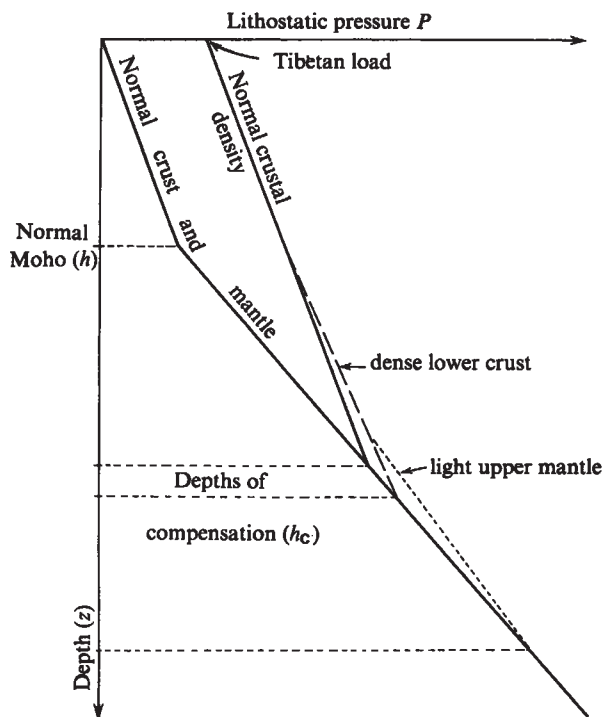


Fig. 6 Schematic plot of lithostatic pressure versus depth under both Tibet and normal crust. Down to normal Moho (h) the difference in lithostatic pressure is approximately constant and proportional to elevation difference $(1.2-1.4) \times 10^8$ Pa. Pressure difference then decreases to zero at depth of compensation h_c . Average value of pressure difference over thickness of layer above depth of compensation is $\sim \sigma_1$, the stress India applies to Eurasia in the plane indentation analogy.

is 5 km higher than the surrounding stable continental areas, for all depths shallower than the depth of compensation, the lithostatic pressure will be greater than that beneath the surrounding stable areas (Fig. 6). If we integrate this pressure difference from the surface to the depth of compensation, we obtain an estimate of the force per unit length, or the average pressure times depth of compensation, acting along the boundaries of Tibet^{14,16}. The value of this quantity depends upon how compensation occurs, but for all models in which compensation occurs primarily by crustal thickening beneath Tibet, the value is $\sim 5 \times 10^6 \text{ Pa km}^{-1}$ (1 bar = 10^5 Pa).

By dividing this number by the thickness of the zone, H , assumed to behave analogously to a rigid-plastic material, we obtain an average estimate of σ_1 . For a 100-km thick zone, $\sigma_1 = 5 \times 10^7 \text{ Pa}$. From the relationships above, the yield stress, τ_Y , would be between $\sim 6 \times 10^6$ and $1.6 \times 10^7 \text{ Pa}$. For a 50-km thick zone, these values should be doubled.

These estimates of σ_1 are only approximate and depend upon the assumption that the elevation of Tibet and the compensating mass deficiency are not maintained by vertical shear stress. Moreover, they represent average pressures, averaged over an assumed thickness (H), in which mechanical properties vary markedly with depth. If the yield stress estimated above was controlled essentially by the strength of the faults in the upper crust, where strain is relieved by earthquakes, the estimated strength of this brittle region would be larger than the values of τ_Y given above.

For a range of thickness for brittle deformation of 20–40 km (ref. 16), the average stresses on faults would be 2.5–5 times larger than those estimated above, that is as small as $1.5 \times 10^7 \text{ Pa}$ or as large as $8 \times 10^7 \text{ Pa}$. The number of assumptions make it impossible to be more precise. Moreover, stress could concentrate in localised regions on faults (asperities) to much higher levels, and the average stress may not reach $1.5 \times 10^7 \text{ Pa}$ for all earthquakes. In any case, the estimated range of average shear stresses on faults is higher than nearly all calculated stress drops of earthquakes (10^4 – 10^7 Pa), but is $< 10^8 \text{ Pa}$.

Summary

When the pattern of faulting in Asia is compared with slip-line fields calculated for various plastic plane strain indentation problems several seemingly unrelated phenomena in Asia appear to have a common cause. India behaves like a rigid die that indents the rest of Eurasia, causing deformation over a large area that resembles deformation in rigid-plastic media. The linear Herat and Altyn Tagh faults are similar to α and β lines, respectively, caused by wedge indentation by north-western India. The more curved Kunlun and Kangting faults

are similar in sense and trend to slip lines near the edge of a more planar indenter, in eastern India. The right lateral Red River fault and the convergence at the Burman arc are parallel to α lines and the direction of plastic flow, respectively, for deformation near the corner of an indenter that has already penetrated deeply into a rigid-plastic solid. The simultaneous convergence at the Himalayas and flow around its eastern end causing convergence at the Burman arc is responsible for the Himalaya–Burman syntaxis. Tension is predicted at a large distance from a flat indenter, and the Baikal Rift Zone and Shansi Graben may be consequences of an analogous state of stress in Asia. From an estimate of the indentation pressure needed to maintain the high elevation of Tibet, we estimate that the average yield stress in a 100-km thick zone in Asia is $\sim 10^7 \text{ Pa}$. If all of the yielding occurs by stress accumulation on faults and subsequent stress drops during earthquakes to a depth of 20–40 km, the average stress on the faults is a few hundred bars and less than a kbar.

We think that the analogy of plastic plane strain with the tectonic processes in Asia supports the contention that most of the tectonics of Asia are caused by the collision of India with Eurasia and provides a unifying explanation for the phenomena occurring there and in continental collisions in general. This type of approach may lead to an even more general understanding of both the geographical distribution of different tectonic styles within continents at given epochs and the succession in time of different tectonic phases in a given continental area.

C. Goetze, M. Mattauer and F. McClintock clarified several points in this paper, and G. Garcia drafted the figures. This research was supported by the NSF, by the Laboratoire de Geologie Structurale and by an Alfred P. Sloan Fellowship.

Received June 28; accepted October 7, 1976.

- 1 Hill, R., *The Mathematical Theory of Plasticity* (Oxford University Press, 1950).
- 2 Backofen, W. A., *Deformation Processing* (Addison-Wesley, Reading, Massachusetts, 1972).
- 3 Johnson, W., Sowerby, R., and Haddow, J. B., *Plane Strain Slip-line Fields: Theory and Bibliography* (American Elsevier Publishing Co., Inc., New York, 1970).
- 4 McKenzie, D. P., *Geophys. J. R. astr. Soc.*, **30**, 109 (1972).
- 5 Molnar, P., and Tapponnier, P., *Science*, **189**, 419 (1975).
- 6 Tapponnier, P., and Molnar, P., *J. geophys. Res.* (in the press).
- 7 Mattauer, M., *Earth planet. Sci. Lett.*, **28**, 144 (1975).
- 8 Mattauer, M., *Les Deformations des Matériaux de l'Encre Terrestre* (Hermann, Paris, 1973).
- 9 Tapponnier, P., and Francheteau, J., *J. geophys. Res.* (in the press).
- 10 Mohajer-Ashjai, A., Behzadi, H., and Berberian, M., *Tectonophysics*, **25**, 281 (1975).
- 11 Tchalenko, J. S., Braud, J., and Berberian, M., *Nature*, **248**, 661 (1974).
- 12 Peire, A. V., Burtman, V. S., Ruzhentsev, S. V., and Suvorov, A. I., *Int. Geol. Cong. 22nd Session, Delhi, Part XI*, 441 (1964).
- 13 Fitch, T. J., *J. geophys. Res.*, **75**, 2699 (1970).
- 14 Frank, F. C., in *Flow and Fracture of Rocks* (Geophysical monograph 16, American Geophysical Union, Washington, DC, 1972).
- 15 McKenzie, D. P., in *The Nature of the Solid Earth*, 323 (McGraw-Hill, New York, 1972).
- 16 Chen, W.-P., and Molnar, P., *J. geophys. Res.* (in the press).

A complete skeleton of the Late Triassic ornithischian *Heterodontosaurus tucki*

Albert P. Santa Luca

Peabody Museum, Harvard University, Cambridge, Massachusetts 02138

A. W. Crompton

Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138

Alan J. Charig

Department of Palaeontology, British Museum (Natural History), London SW7 5BD, UK

The first ornithischian dinosaur to be found in the Triassic (other than very fragmentary material) was Heterodontosaurus tucki from South Africa, described¹ on its skull. The postcranial skeleton of another H. Tucki is described here. Heterodontosaurus is compared with its sympatric contemporary Fabrosaurus, the only other well known Triassic ornithischian, and the interrelationships of early dinosaurs are discussed briefly.

LATE Triassic ornithischian dinosaurs fall into two distinct families: the Heterodontosauridae, including the genera *Heterodontosaurus*², *Lycorhinus*², *Abriotosaurus*² and *Lanasaurus*³, and the Fabrosauridae with the single genus *Fabrosaurus*⁴⁻⁶. (*Fabrosaurus*, however, is considered by some⁷ to be gen. indet.) The general structure of the skull in these two families is known in broad detail^{1,4,5,7-9}, and in *Fabrosaurus* the postcranial skeleton too is fairly well known⁶. Very little has been published hitherto, however, on the postcranial skeleton of *Heterodontosaurus*. This article is essentially a brief account of a virtually complete skeleton of *H. tucki*.

The specimen (South African Museum no. K1332) was discovered in December 1966 in the Upper Red Beds of the Stormberg Series, at a height of about 1,770 m above sea level on the northern slopes of Krommespruit Mountain in the District of Herschel, Republic of South Africa. More complete locality data will be published when the skull and skeleton are described in detail. The scapulo-coracoid, humerus and hand of the specimen have already been partly figured elsewhere¹⁰, though to a very small scale. The only other Triassic ornithischian known with postcranial material worth mentioning is the poorly preserved and fragmented *Pisanosaurus mertii* (ref. 11 and J. F. Bonaparte, unpublished) from the Ischigualasto Formation of Argentina.

are transitional in form and the exact position of the boundary between neck and trunk cannot be determined objectively; however, 9 cervicals plus 12 dorsals seems to be the most reasonable allocation.) Twenty-eight caudal vertebrae are preserved in two continuous series, nos 1-11 and (probably) nos 17-33, separated by a gap representing about five vertebrae. It is not possible to estimate accurately the number of vertebrae required to complete the column to the tip of the tail, but it could not have been many because the last vertebrae preserved are very small. Chevrons are present between and below most of the caudals. The remains of ossified tendons lie alongside the neural arches from the middle of the trunk to the posterior end of the sacral series, but they are absent from the tail.

Pectoral girdle and forelimb

The scapula and coracoid are slightly curved, with the concave surface medially; the curve of the coracoid is a little more pronounced than that of the scapula. The dorsal region of the scapula forms a wide blade and appears to have been capped by a cartilaginous extension, while below the blade the scapula narrows to a nearly circular cross section. The wider ventral region supports a well developed acromial process on its anterior margin. A powerful tubercle, presumably for the origin of the scapular head of the triceps muscle, is situated immediately

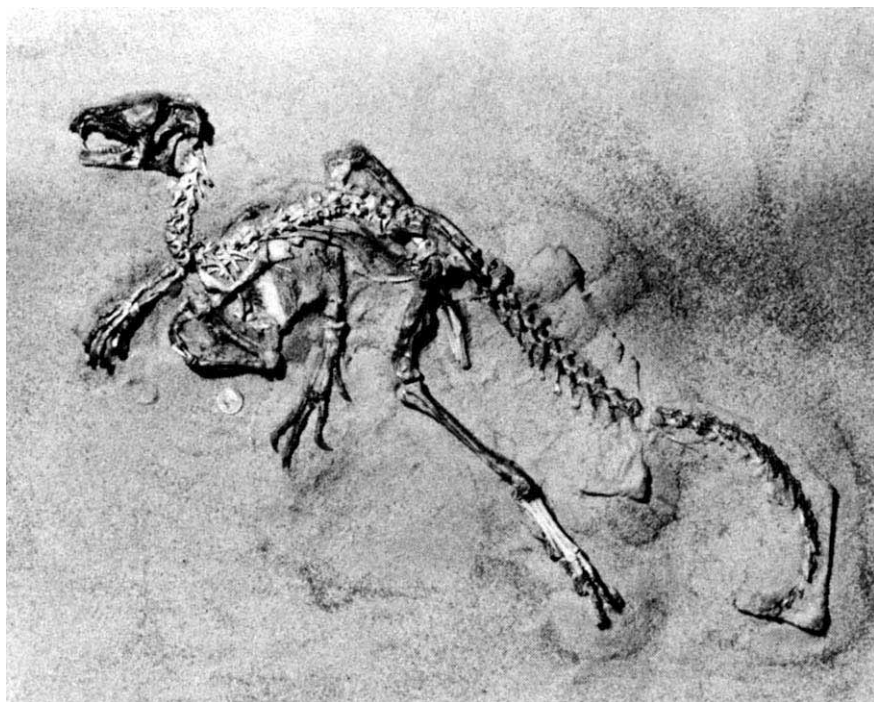


Fig. 1 *Heterodontosaurus tucki*. The specimen described in the text (SAM no. K1332).

The skeleton of *Heterodontosaurus* described here is preserved in an articulated condition with little displacement or distortion (Fig. 1). Figure 2 is the first ever complete reconstruction of the skeleton of a Triassic ornithischian. Measurements of some of the elements are given in Table 1.

Axial skeleton

The skull is almost identical with that of the holotype^{1,7,8}. The vertebral column includes 21 presacral vertebrae and 6 sacrals, the latter fused together. (These figures, 21+6, compared with the typical thecodontian count of 25+2, suggest that the four additional vertebrae incorporated into the sacrum are all dorsals rather than caudals. The presacral count of 21 has not been broken down into separate cervical and dorsal counts because presacrals 8-10

above the dorsal lip of the glenoid. Near the forelimb lies a roughly rectangular sternal plate with a weakly concave anterior edge, a convex medial edge and straight lateral and posterior edges.

The humerus is slightly twisted along its longitudinal axis; its head is distinct, and the strongly developed deltopectoral crest is so long that it occupies about 40% of the length of the whole bone. There is no olecranon fossa on the posterior surface, in spite of the presence of a distinct olecranon process on the ulna; the anterior surface of the humerus, however, bears a shallow depression above the condyle. The entepicondyle is large, suggesting that the flexor muscles for the forearm and manus were strong. The radius too is slightly twisted about its longitudinal axis, and a small crest, 10 mm high, arises from the middle of the lateral surface of the shaft. The carpus consists of nine

bones: a radiale, an ulnare, a pisiform, a centrale and five distal carpals. The ulnare, with a highly concave proximal surface and a nearly flat distal surface, is the largest of these; a small foramen pierces its dorsal surface. The radiale has the same transverse width as the ulnare but is flatter. Of the distal carpals only V shows reduction, in spite of the fact that digits IV and V are both reduced; indeed, distal carpal V is actually smaller than the pisiform.

The manus is extremely long, the length of digit III being nearly 52% of that of the humerus. The phalangeal formula is 2-3-4-3-2. The distal articular surface of each of the metacarpals I-III is extended dorsally and proximally on to the dorsal surface to an unusual degree; this suggests that the proximal phalanx of the corresponding digit could have been hyperextended on the metacarpal. The distal end of metacarpal I is twisted relative to its proximal end, likewise the distal end of the proximal phalanx of both digits II and III; this would have caused digits I-III to move mediad during flexion, thus producing an effective grasping motion of the manus. Further, the terminal claw-bearing phalanges (unguals) on those three digits bear pronounced flexor tubercles.

Pelvic girdle and hindlimb

The ilium has a long, shallow anterior process which terminates in a small knob-like expansion. The posterior process is slightly shorter and slightly deeper and its posterior

marks the attachment of the ischio-trochanteric and internal tibial flexor muscles¹⁴.

The femur is relatively short and slightly curved with a convex anterior surface; the head projects nearly at right angles to the longitudinal axis of the shaft, the greater trochanter is not separated from the lesser trochanter by a marked cleft, and a pendent fourth trochanter, about 10 mm long, forms an angle of approximately 45° with the axis of the shaft. The tibia is very long, 1.29 times the length of the femur; its proximal articulating surface is not horizontal but slopes downwards and outwards from its medial edge. The proximal part of the tibia is considerably expanded anteriorly into a strong cnemial crest. The head of the fibula also is expanded antero-posteriorly but, even so, its length is only half that of the tibial head; distally the fibula narrows to a thin rod which fuses with the tibia. The astragalus and the calcaneum are indistinguishably fused with the tibia and fibula. The three distal tarsals are fused with the metatarsal heads; thus the lateral distal tarsal caps metatarsal IV, the central caps III, and the medial distal tarsal covers both II and I.

The first four metatarsals are fused to form a single unit, their proportional lengths (taking the sum of their lengths as 100%) being 16.8%, 26.1%, 30.0% and 27.1% respectively. Metatarsal V, however, is reduced to a slender splint. The phalangeal formula is 2-3-4-5-0. Digit III continues the longitudinal axis of its metatarsal, but digit II

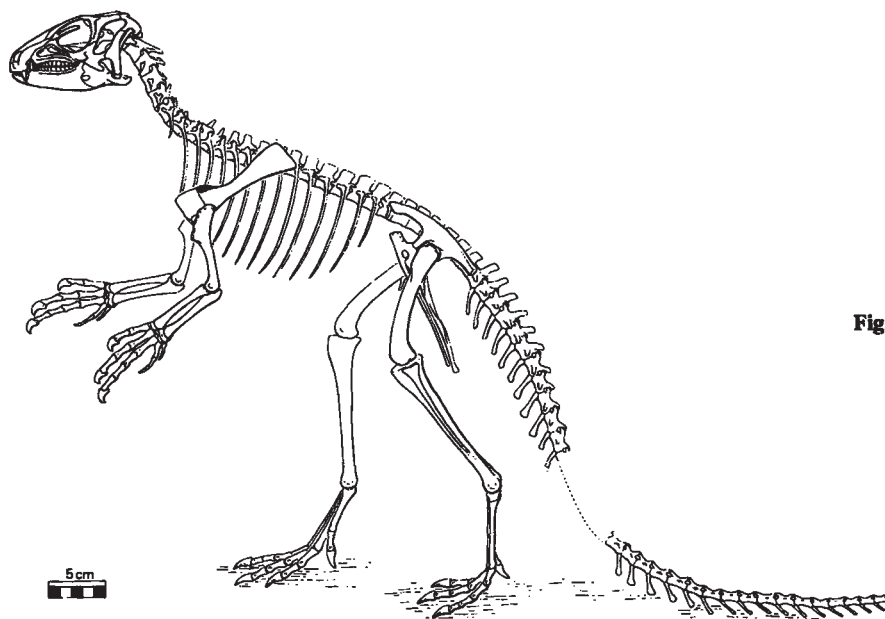


Fig. 2 *Heterodontosaurus tucki*. Reconstruction of the complete skeleton.

end also is a little expanded. The pubic peduncle lies almost at right angles to the longitudinal axis of the ilium. A mound-like protuberance and an articular facet on the lateral surface of the ischiadic peduncle increase the depth of the acetabulum. The pubis has a short, deep anterior ramus (15 mm long) and a long, straight, rod-like posterior ramus (pubis proper) which is nearly as long as the ischiadic rod lying immediately above it. Several small rounded tubercles are present on the lateral surface of the anterior ramus; these may be associated with the origins of the ambiens muscle and of the three heads of the pubo-ischio-femoralis internus¹²⁻¹⁴. The ischium is joined to the ilium by a short rounded process, but the process which attaches it to the pubis is flat, wide and expanded. Posterior to this pubic process the ischium narrows to a straight slender rod with a slightly expanded end; there is no obturator process on the ischium, but a prominent and gently curved crest on the central part of the rod probably

diverges medially and IV laterally; digit I is partly reduced and apparently had no role in support during locomotion. The unguals are slightly curved.

Discussion

Heterodontosaurus tucki seems to have been a bipedal cursorial herbivore; this opinion is based mainly on the structure of the manus and pes, the length ratio of tibia to femur, and the structure of the teeth and jaws. Although the forelimb may have been used for locomotion when the animal was moving slowly, for example when foraging, it seems to have been adapted primarily for grasping and tearing. A large olecranon process is usually present in quadrupeds like *Stegosaurus* but absent in bipeds such as *Fabrosaurus*; its presence in *Heterodontosaurus tucki*, however, may have been functionally associated with the powerful arm and highly manipulative hand, for the

Triassic coelurosaur *Syntarsus*¹⁵ also has both an olecranon process and a grasping hand.

The structure of the postcranial skeleton supports the view, based originally upon the skull structure, that the two families of Late Triassic ornithischian dinosaurs (namely the Fabrosauridae and the Heterodontosauridae) differ from one another in many important features. (This, in turn, accords with the belief that each family had a long independent history dating back to at least mid-Triassic times.) *Fabrosaurus* too was a cursorial biped, sympatric with *Heterodontosaurus*, but it lacks the deep skull of *Heterodontosaurus* and its powerful and specialised masticatory apparatus. Only five sacral vertebrae are present in *Fabrosaurus* instead of six. The deltopectoral crest and whole forelimb are relatively shorter in *Fabrosaurus* than in *Heterodontosaurus*, and *Fabrosaurus* does not have a long and powerful grasping hand. Further, the ulna of

Table 1 Selected measurements* of *Heterodontosaurus tucki* (SAM no. K1332)

	(mm)
Scapula, maximum length	86.9
Coracoid, maximum length	22.3 estimated
Humerus, maximum length	82.3
Radius, maximum length	58.5 (right)
Ulna, maximum length	67.6
Metacarpals, maximum lengths: I	17.6
II	23.3 (right)
III	22.4
IV	14.5 (right)
Ilium, maximum length	96.9
minimum height above acetabulum	15.0
Pubis, maximum length	125.7 (right)
height of anterior ramus	7.8 (right)
Ischium, maximum length	114.2 (right)
Femur, maximum length	112.2
length of fourth trochanter	9.8
Tibia, maximum length	145.0 (right)
Metatarsals, maximum lengths: I	38.1 (right)
II	59.1 (right)
III	67.9 (right)
IV	61.4 (right)
Combined length of centra of twelve dorsal vertebrae	171.7
Hindlimb (femur + tibia + metatarsal III)	1.89
Trunk (12 dorsal centra)	

* Left side unless stated otherwise.

Fabrosaurus lacks the olecranon process and its humerus lacks the prominent entepicondyle. The posterior process of the ilium is shorter and deeper, the pubic peduncle is directed forwards and downwards relative to the long axis of the iliac blade (rather than nearly at right angles thereto as in *Heterodontosaurus*) and there is a large obturator process on the ischium. There are also many differences between the two genera in the structure of the hindlimb. In *Fabrosaurus* the greater trochanter is separated from the lesser trochanter by a deep cleft, the fibula has an expanded distal end and the astragalus and calcaneum are fused neither with the tibia and fibula nor with each other. There is, however, one important feature of the pelvis common to both *Fabrosaurus* and *Heterodontosaurus*: a complete and undescribed pubis of *Fabrosaurus* (SAM no. K1106) shows that the anterior ramus is very similar in shape and even in size to that of *Heterodontosaurus*.

The skeletal remains of *Pisanosaurus*¹¹ are very incomplete, but the lower jaw and dentition suggest a closer affinity with *Heterodontosaurus* than with *Fabrosaurus* (Ref. 7 and J. Bonaparte, unpublished). The separate fibula of *Pisanosaurus* with its unreduced distal end and the astragalus and calcaneum fused neither with the tibia and fibula nor with each other might be thought to suggest a relationship to *Fabrosaurus* closer than to *Heterodontosaurus*, but in fact such features are merely primitive and without phylogenetic significance.

The combination of specialised features seen in *Heterodontosaurus* seems to rule it out as a possible ancestor of known ornithopod ornithischians of later times¹⁶. These features include the nature of the dentition (the closely set teeth worn down to a continuous occlusal surface), the low number of dorsal vertebrae, the large grasping hand, the absence of the obturator process on the ischium, the reduced fibula and the complete fusion of the distal ends of the tibia and fibula with each other and with the proximal tarsals. Although so early in time, *Heterodontosaurus* is far more specialised in these and other characters than are the majority of much later ornithopods—such as *Hypsilophodon* and *Iguanodon* of the Early Cretaceous and even *Thescelosaurus* of the Late Cretaceous. Possible relationships can be discussed in detail only when the available material of Triassic dinosaurs has been more fully prepared and studied.

Bakker and Galton¹⁰ have stated that all Triassic dinosaurs share several common features, such as the shape and direction of the glenoid cavity, the position of the deltopectoral crest, the structure of the hand, the perforation of the acetabulum, the structure of the femur (in particular the position and size of the fourth trochanter), the restriction of femoral movements to the vertical plane, and the mesotarsal type of ankle joint. On the basis of these similarities they have concluded that the three major groups of Triassic dinosaurs (prosauropods, theropods and ornithischians—the last in effect restricted to ornithopods) arose from a central stock of early forms such as *Staurikosaurus*¹⁷, *Ischisaurus*¹⁸ and *Herrerasaurus*^{18,19}. This stock, according to them, could have been derived in turn from a group of thecodontians of which *Lagosuchus*²⁰ is a representative. Such a conclusion, however, seems hardly to accord with the view that even the Triassic ornithischians alone comprise two distinct families with long independent histories, probably going back to no later than mid-Triassic times. A fundamental distinction between the three main groups of Triassic dinosaurs lies in the structure of the pelvic girdle; the form of the prosauropod pelvis does not differ greatly from the primitive thecodontian pattern, whereas the theropod and ornithischian pelves are more highly modified and are also very different from each other. Bonaparte²⁰ is not totally in agreement with Bakker and Galton¹⁰ and still believes (following Charig^{21,22}, Charig *et al.*²³) that both saurischian suborders, as well as the ornithischians, were derived independently from thecodontians. It may well be that some of the common characteristics of Triassic dinosaurs as listed by Bakker and Galton were present in an ancestral thecodontian stock and that others were the result, at least in part, of parallel evolution. Some features of the hand of *Heterodontosaurus* seem to support this view; for example, the facts that metacarpals IV and V of *Heterodontosaurus* are divergent and that digit I does not diverge markedly from digit II are more reminiscent of the hands of thecodontians than of those of early saurischians. Likewise, one of us argues²⁴ that most of the alleged similarities between Saurischia and Ornithischia do not exist or are without phylogenetic significance; the lack of common characters 'primitive' to the supposed taxon Dinosauria, together with the existence of several important differences (not mentioned by Bakker and Galton) between the two orders, suggests that the alleged monophyly of the "Dinosauria" cannot be demonstrated.

It seems undeniable that there was a unity to dinosaur biology which greatly exceeded a few osteological features, but this may well have been due to the fact that all the groups concerned were in any case fairly closely related and had been subjected to the same evolutionary pressures. Whatever view is taken of the controversy concerning the origins and interrelationships of the major groups of dinosaurs, all will agree that it calls for the thorough and detailed preparation and description of the now extensive

collections of Triassic ornithischian dinosaurs and a comparison of these with all available material of their saurischian contemporaries.

We are indebted to Dr T. H. Barry, Director of the South African Museum (Cape Town), for permission to study the specimen described in this article.

Received September 17; accepted October 1, 1976.

¹ Crompton, A. W., and Charig, A. J., *Nature*, **196**, 1074–1077 (1962).

² Hopson, J. A., *S. Afr. J. Sci.*, **71**, 302–305 (1975).

³ Gow, C. E., *Zool. J. Linn. Soc.*, **57**, 335–339 (1975).

⁴ Thulborn, R. A., *Palaeontology*, **13**, 414–432 (1970).

⁵ Thulborn, R. A., *J. Zool., Lond.*, **164**, 165–179 (1971).

⁶ Thulborn, R. A., *Palaeontology*, **15**, 29–60 (1972).

⁷ Charig, A. J., and Crompton, A. W., *Ann. S. Afr. Mus.*, **64**, 167–189 (1974).

⁸ Charig, A. J., and Crompton, A. W., *Ann. S. Afr. Mus.* (in the press).

⁹ Thulborn, R. A., *Zool. J. Linn. Soc.*, **55**, 151–175 (1974).

¹⁰ Bakker, R. T., and Galton, P. M., *Nature*, **248**, 168–172 (1974).

¹¹ Casamiquela, R. M., *Ameghiniana*, **5**, 47–64 (1967).

¹² Romer, A. S., *J. Morph.*, **43**, 347–385 (1927).

¹³ Romer, A. S., *Acta zool., Stockh.*, **8**, 225–275 (1927).

¹⁴ Galton, P. M., *Postilla*, **131**, 1–64 (1969).

¹⁵ Raath, M. A., *Arnoldia (Rhodesia)*, **4**, 1–25 (1969).

¹⁶ Galton, P. M., *Bull. Br. Mus. nat. Hist. (Geol.)*, **25**, no. 1, 1–152 (1974).

¹⁷ Colbert, E. H., *Am. Mus. Novit.*, No. 2405, 1–39 (1970).

¹⁸ Reig, O. A., *Ameghiniana*, **3**, 3–20 (1963).

¹⁹ Benedetto, J. L., *Ameghiniana*, **10**, 89–102 (1973).

²⁰ Bonaparte, J. F., *Acta geol. Iliana*, **13**, 5–90 (1975).

²¹ Charig, A. J., *Diss., Univ. Cambridge* (1956).

²² Charig, A. J., *Abstr. Diss. Univ. Camb.*, 1955–56, 28–29 (1957).

²³ Charig, A. J., Attridge, J., and Crompton, A. W., *Proc. Linn. Soc. Lond.*, **176**, 197–221 (1965).

²⁴ Charig, A. J., in *Morphology and Biology of Reptiles* (edit. by Bellairs, A. d'A., and Cox, C. B.), 65–104 (Academic, London, 1976).

Cell length, cell growth and cell division

W. D. Donachie, K. J. Begg & M. Vicente

Department of Molecular Biology, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

When cells of E.coli reach a certain critical length, which is constant in all growth conditions and equal to twice the minimum cell length, they abruptly increase their rate of elongation and divide about 20 min later. Chromosome replication terminates at about this same cell length but is not the signal for the change in rate of cell elongation.

DURING exponential growth, cells of *Escherichia coli* divide after each doubling in their volume. Their growth rate, and thus the frequency at which they divide, depends on the temperature and the availability of nutrients in the culture medium. *E.coli* cells are cylinders with hemispherical ends and, at any one growth rate, the doubling in cell volume between successive division takes place entirely by a doubling in cell length, without detectable change in cell diameter¹. But at each growth rate one observes a different cell volume, length and diameter. The relationship between cell volume and growth rate has been known for some time² and explanations for this relationship have been given^{3–5}. Nevertheless there has not, to our knowledge, been a comparable systematic investigation of the relationship between growth rate and cell length in *E.coli* (although there are several scattered measurements in the literature). We have therefore carried out such an investigation and have found a simple relationship between cell length, cell growth and cell division. This is that individual cells become committed to division and increase their rate of elongation at a fixed length which is independent of growth rate. Moreover, the change in rate of elongation seems to depend solely on the attainment of the critical length and not on other factors such as the termination of chromosome replication (as has been suggested by others^{6–8}). Following a change in growth conditions, cells do not alter their rate of elongation so that it is proportional to their new rate of volume and mass increase, until they attain the critical length. As a consequence, cells change their size and proportions to those characteristic of the new growth rate within one division cycle after the change in growth conditions.

These experimental results are discussed in terms of a modified "unit cell" model⁹ in which the basic unit of cell growth is a section of cell of fixed length, Λ , which extends in length at a fixed number of growth sites (possibly one). The rate of cell elongation therefore doubles after completion of a new unit cell, at length 2Λ . The completion of a new unit cell occurs at the same time as does a signal for the commencement of septation. Septation takes approximately 20 min to complete in all growth conditions.

Relationship between cell length and growth rate

We have calculated mean cell lengths from cell length distributions of *E.coli* B/r (ATCC 12407) growing exponentially at

37 °C in a number of different media (Table 1). A linear regression line was calculated using the method of least squares, which gave

$$\bar{L} = (0.67R + 2.01) \mu\text{m} \quad (1)$$

(where $\bar{L}$ is mean cell length in μm and R is the number of doublings in cell number h^{-1}). The correlation coefficient between the data and the regression line was $r = 0.96$ (Fig. 1).

Average cell lengths were also calculated for a K12 strain, OV2 (Table 1) growing in different media at three different temperatures. These points are also shown in Fig. 1, from which it can be seen that they obey the same relationship as the unrelated B/r. (For these points, $\bar{L} = 0.6R + 2.03$; for which $r = 0.88$.)

The relationship between mean cell length and growth rate shown in Fig. 1 has the following implications: (1) Cell length at a time approximately 20 min before division is approximately the same at all growth rates. (2) This length is approximately twice the minimum length of a cell in any growth condition.

We reach these conclusions as follows. For this argument we rewrite equation (1), without significant error, as

$$\bar{L} = (2 + 2R/3) \mu\text{m} \quad (2)$$

The length of cells at birth (L_B) and at division (L_D) may be calculated, to a first approximation, as proportional to cell age, using the theoretical distribution of cell ages in an exponential population^{1,10}, so that

$$L_B = \bar{L} \ln 2 \quad (3)$$

$$\text{and } L_D = 2L_B \quad (4)$$

(Mathematically more sophisticated methods of estimating these lengths give very similar results when applied to our population; see ref. 12.) Equations (2), (3) and (4) therefore allow us to calculate the lengths of cells at division for all growth rates. But to calculate cell lengths at different stages in the cell cycle, we must make some assumption as to how the cells increase in length (for example, at a constant rate or at a rate which increases as the cell cycle progresses). It is known that cells of *E.coli* increase in length continually over the cell cycle^{1,9,13} but the exact shape of the curve of increase is difficult to determine experimentally, both because of the smallness of the cells and because the change in rate of elongation is only twofold. Nevertheless it can be assumed that this curve lies somewhere between the extreme of a straight line (that is, a constant rate of elongation throughout the cycle) and a linear function in which the rate doubles some time during the cycle.

(An exponential rate of elongation would lie between these two curves.) In the special case in which the rate of increase is linear with the doubling 20 min before division, equations (2), (3) and (4) show that cell length at this time will be the same for all growth rates (Fig. 2b) and that this length is

$$(4 \ln 2) \mu\text{m} = 2.77 \mu\text{m}$$

Even in the cases where length increases exponentially (Fig. 2a) or at a constant rate over the whole cycle, these equations show that cell length 20 min before division will be approximately the same at all growth rates. These lengths will be $2.86 \mu\text{m} \pm 3\%$ and $2.96 \mu\text{m} \pm 6\%$ respectively. Thus cell length must be approximately $2.8 \mu\text{m}$ at a point 20 min before division at all growth rates.

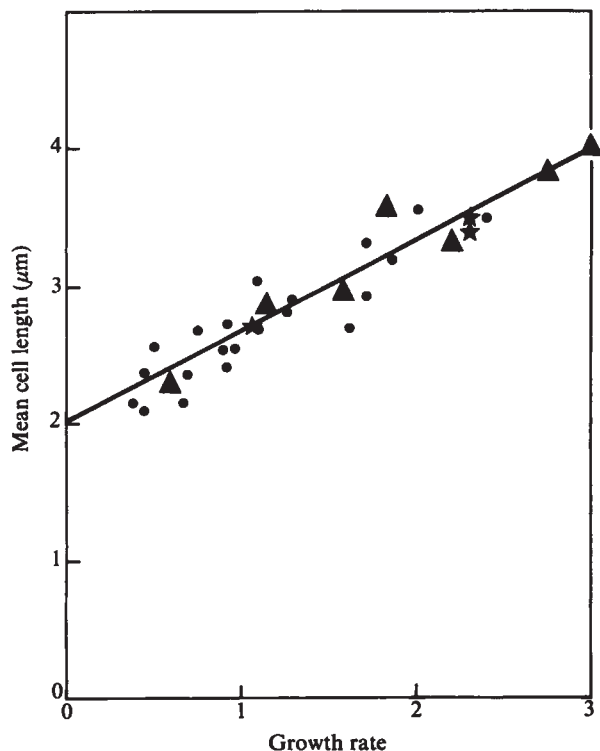


Fig. 1 Mean cell length of exponential phase populations of *E. coli* (data from Table 1) as a function of growth rate (doublings in cell number per hour). The line is the best-fit linear regression calculated by the method of least squares. (Triangles, B/r ATCC 12407; circles, OV2; measurements by light microscopy. Stars, OV2; measurements by electron microscopy).

The minimum length of cells in any steadily growing population may be estimated as the birth length at zero growth rate, using equations (2) and (3). The calculated value for this theoretical minimum length, Λ , is

$$\Lambda = 2 \ln 2 \mu\text{m} = 1.39 \mu\text{m} \quad (5)$$

Thus cell length 20 min before division must always be approximately twice the theoretical minimum cell length. In the special case shown in Fig. 2b, this length must be precisely 2Λ at all growth rates.

In the various models for cell elongation discussed above, it has been assumed that the rate of elongation will be a function of the growth rate in mass and number. This assumption works well for steady growth in a constant environment but it can easily be seen that it cannot apply without modification to cell growth during the transition between one growth rate and another. Thus, taking the model shown in Fig. 2b as an example, if the rate of elongation of a cell were to change immediately to that characteristic of the new environment after a shift from one growth medium to, say, a richer one, cell length at the time of

the next division would be greater than that actually found in the new medium. In fact, if the rate of elongation did change immediately after such a shift, cells would possibly maintain a constant diameter at all growth rates. This is not so. One possible solution to this paradox would be if the rate of cell extension remained the same after a shift to a new mass doubling time and changed to that characteristic of the new growth rate only after the cells reached a length of 2Λ . Our observations support this prediction and, in so doing, also provide evidence that the rate of elongation does change abruptly at about the time and cell length predicted by the model shown in Fig. 2b.

Figure 3a shows such an experiment. At zero time, a selected fraction of small cells was transferred from a poor growth medium (generation time 67 min) to a rich medium (generation time 25 min). The theoretical time of first division after such a transfer is about 60 min (ref. 14) and cells in fact divided between 40 and 60 min (Fig. 3a). The rate of mass increase of the cells increased immediately to the new rate¹⁵ but the rate of elongation remained at the calculated pre-shift rate until it increased abruptly by 5.7-fold after about 45 min. (Figure 3b shows a control experiment in which the selected cells were maintained in the same poor medium as before selection. In this case, the point at which the rate of elongation increased could not be determined, as expected, but the increase in rate over the first division cycle was close to twofold.)

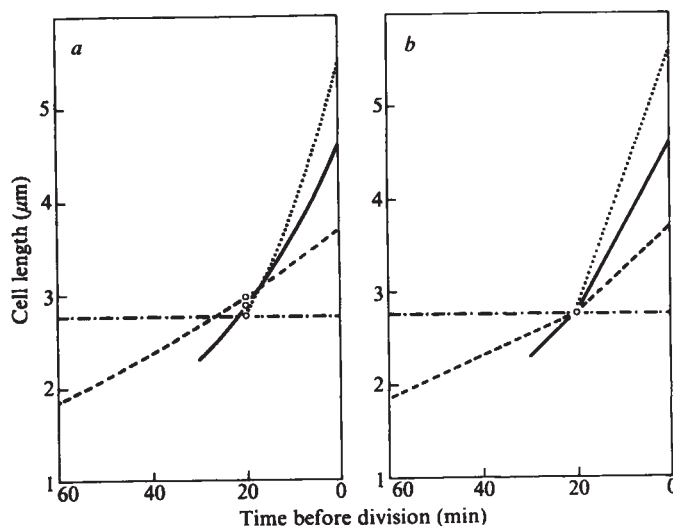


Fig. 2 Calculated course of increase in cell length over cell cycles of various durations, assuming that length doubles over the course of each cycle, that the relationship between mean population cell length and growth rate is given by the straight-line function (equation (2), Fig. 1), and that populations have the theoretical age distribution¹⁰ (see text). The figure shows cell elongation according to two sets of assumptions: *a*, that the rate of elongation increases continuously according to a logarithmic function, and *b*, that the rate of elongation is constant but doubles 20 min before division. For clarity, only four growth rates are shown, namely those for generation times of 20 min (---), 30 min (—), 60 min (— — —) and infinity (— · —). The open circles represent cell length at 20 min before division.

The model we propose to describe cell elongation is therefore that newborn cells elongate at a constant rate which is proportional to the rate of growth in mass. Such newborn cells will always be between Λ and 2Λ in length and will reach 2Λ after a period of growth which is proportional to the initial cell length and the growth rate. In steady growth conditions, the rate of elongation will double at this point. (During a medium shift, this rate increase will be a doubling times the change in growth rate in mass.)

(The possibility that the sucrose gradient selection might perturb the rate of cell elongation may be discounted, since we observe the change in rate of cell extension at the same cell

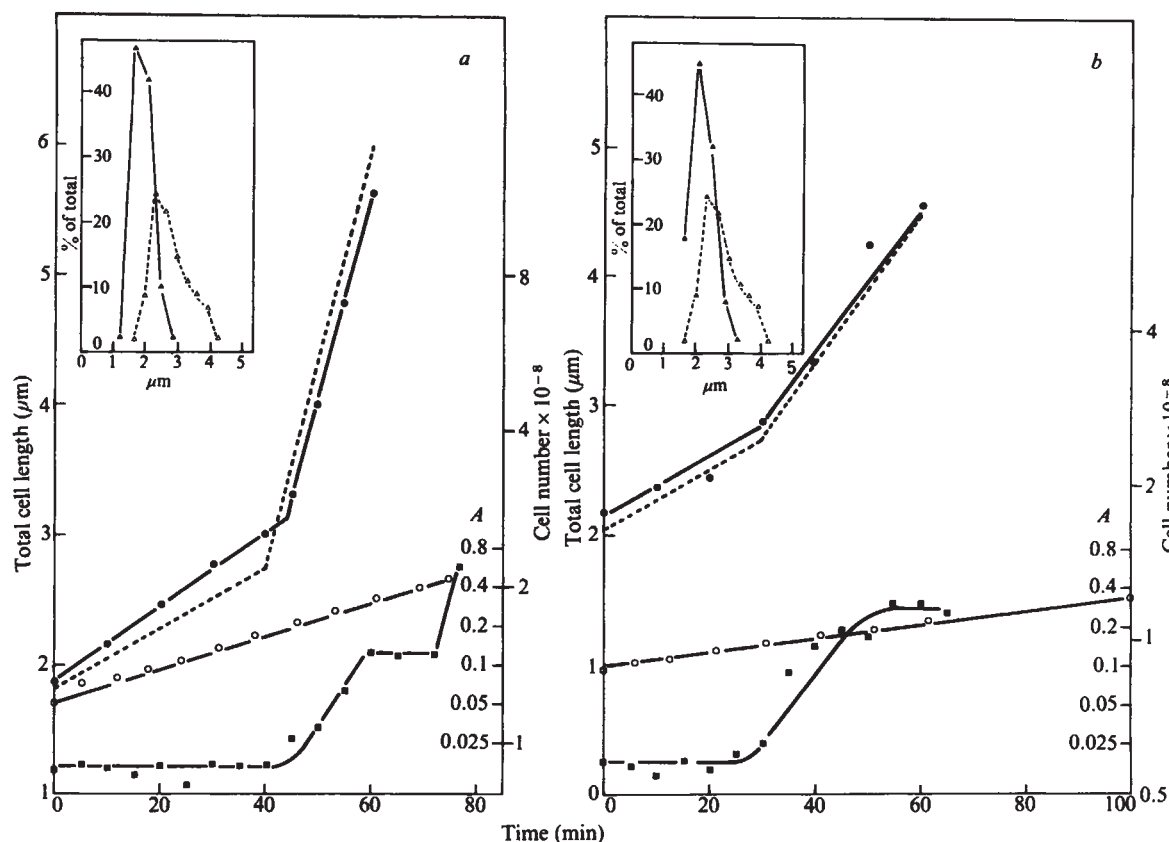


Fig. 3 Course of increase of total cell length (●), total cell mass (○) and total cell number (■) in synchronous cultures of *E. coli* B/r ATCC 12407. (Total cell length is calculated as mean cell length × cell number per cell number at time zero. The error of estimate of mean cell length is $\pm 3\%$). In *a*, small cells were selected by sucrose gradient centrifugation²⁴ from a population growing exponentially in medium 2 (Table 1) with a generation time of 67 min and reinoculated into L-broth + glucose, after which the mass doubling time immediately decreased to 25 min. Cell division took place between about 40 and 60 min. Cell length increased at a constant rate for approximately 40 min and then increased abruptly to a new rate which was about $5.7\times$ the initial rate. In consequence, the cell length and cell mass changed from those for cells growing at a rate of 0.9 doublings h^{-1} to those for cells growing at 2.4 doublings h^{-1} by the time of the first division in the new medium. The predicted course of cell elongation according to the theory developed in this paper is also shown (—). *b*, The result of a control experiment in which small cells from an exponential population with a generation time of 63 min (medium 2) were reinoculated into a fresh batch of the same medium. The selected fraction of cells were slightly larger than those obtained in *A* and there was a small initial stimulation of the rate of mass increase. The theoretical curve for length increase was drawn taking these factors into account. In this case the increase in rate took place at approximately 30 min and was an increase of about $2.3\times$, as predicted. The insets show the length distributions of the selected populations at zero time, superimposed on the distribution for exponential asynchronous populations in the same growth conditions.

Table 1 Growth rate and cell length

Strain	Medium	<i>T</i>	37 °C		<i>T</i>	30 °C		<i>T</i>	42 °C	
			<i>R</i>	$\bar{L}$		<i>R</i>	$\bar{L}$		<i>R</i>	$\bar{L}$
B/r ATCC	1	100	0.60	2.29						
B/r ATCC	2	52.5	1.14	2.88						
B/r ATCC	3	38	1.58	2.96						
B/r ATCC	4	33	1.82	3.54						
B/r ATCC	6	27	2.22	3.03						
B/r ATCC	7	22	2.73	3.83						
B/r ATCC	9	20	3.00	3.99						
OV2	1	67	0.90	2.54	135	0.44	2.09	135	0.44	2.37
OV2	2	65	0.92	2.73	160	0.38	2.13	120	0.50	2.57
OV2	3	65	0.92	2.41	87	0.69	2.36	90	0.67	2.15
OV2	4	47.5	1.26	2.82	80	0.75	2.71	62.5	0.96	2.55
OV2	5	35	1.71	2.93	57	1.08	2.70	—	—	—
OV2	6	—	—	—	—	—	—	37	1.62	2.68
OV2	7	35	1.71	3.32	55	1.09	3.06	32.5	1.85	3.19
OV2	8	30	2.00	3.56	47	1.28	2.89	25	2.40	3.49

OV2 is a spontaneous *ilv⁻his⁻* derivative of *E. coli* K12 M6K (see Table 2). *T* is the cell number doubling time in minutes, *R* is the number of doublings per hour, and $\bar{L}$ is the mean cell length of the population in μm . Media: (1) M9+sodium succinate (0.4%); (2) M9+glycerol (0.8%); (3) M9+glucose (0.4%); (4) M9+glucose (0.4%)+histidine ($20\mu g\ ml^{-1}$)+methionine ($20\mu g\ ml^{-1}$); (5) M9+glucose (0.4%)+casamino acids (0.5%); (6) M9+glucose (0.4%)+casamino acids (0.5%)+tryptophan ($20\mu g\ ml^{-1}$); (7) M9+glucose (0.4%)+casamino acids (0.5%)+tryptophan ($20\mu g\ ml^{-1}$)+cytosine ($20\mu g\ ml^{-1}$)+adenine ($20\mu g\ ml^{-1}$)+uracil ($20\mu g\ ml^{-1}$); (8) Oxoid Nutrient Broth no. 2; (9) Oxoid Nutrient Broth no. 2+glucose (0.4%). All OV2 cultures contained in addition all required auxotrophic supplements (final concentration of thymine was $50\mu g\ ml^{-1}$). For these measurements, cells were grown in log phase with shaking in liquid medium for several generations and samples taken at a density of $10^8\ ml^{-1}$ (media 1-7) or $3\times 10^7\ ml^{-1}$ (media 8 and 9) on to slides coated with a thin layer of agar⁸ containing 0.1% sodium azide to prevent further cell growth. Cells were photographed with a Zeiss Ultraphot microscope under phase-contrast optics. Enlarged projections of the negatives were measured⁹. At least 200 cells were measured for each distribution. The error of estimate of $\bar{L}$ was not more than 3%.

length in single cells growing in agar following a shift from a poor to a rich medium.)

Control of rate of cell elongation

We now ask what event in the cell cycle is responsible for the observed increase in rate of elongation which takes place about 20 min before division. It is known that the cell becomes committed to divide (even in the absence of further RNA, protein or DNA synthesis) at approximately 20 min before division¹⁶⁻¹⁸. This transition point in the *E. coli* cell cycle coincides with (1) termination of the oldest round of chromosome replication¹⁹, (2) segregation of the nuclear bodies¹², (3) completion of the period of protein synthesis required for division¹⁶, (4) initiation of termination protein synthesis²⁰, (5) onset of sensitivity to penicillin (activation of autolysins)²¹ and (6) attainment of a particular length. (The first visible signs of septation can be detected within a few minutes of this time²².) One or more of these events might therefore be necessary for the increase in rate of elongation to take place.

It has already been suggested^{6,7} that termination of each round of chromosome replication might be a signal for a doubling in the rate of cell elongation. This suggestion stemmed from observations on the relationship between cell volume, length and radius as functions of the transit time, *C*, taken by replication forks to travel along the chromosome. According to this model the rate of cell elongation would be determined solely by (1) the number of chromosome termini per cell, and (2) the doubling time for cell mass.

To test this hypothesis, we investigated whether an increase in the rate of cell elongation would take place in the absence of chromosome termination. Step-up experiments, such as that

described, provide a sensitive technique for such a determination. We have therefore shifted cells from a poor to a rich medium in conditions in which chromosome termination was prevented. We found that this does not prevent the increase in rate of cell elongation at the same time as in controls.

For this experiment we used a low thymine-requiring strain, B/r/1 T⁻ (ref. 9) which has the advantage that it is able to increase in mass at the control rate for nearly 1 h after removal of thymine, although DNA synthesis stops immediately. (Thymine auxotrophs of the other strains mentioned here could not maintain this rate for so long. Nevertheless, we have observed that their lengths increase in a similar way to B/r/1 T⁻ during this time.) It should be noted that this strain is about 50% longer than the others described here.

Small cells were selected as before from an exponential population in minimal medium plus thymine (generation time 56 min) and transferred to a rich medium plus or minus thymine. After transfer both aliquots grew initially with a mass doubling time of 33 min (Fig. 4), although growth in the thymine-starved aliquot slowed down by about 50 min after transfer. Synchronous division of the control cells took place between 35 and 45 min while the thymine-starved cells showed only a small increase in number. The initial rate of length increase in both cultures was initially that appropriate to the pre-shift growth rate but changed abruptly to the expected higher rate (roughly 3.8 times greater) at about 30 min in both cultures. We consider that this experiment shows that termination of chromosome replication is not required as signal for the observed abrupt increases in rate of cell elongation.

This conclusion is further corroborated by the observation that an increase in the number of chromosome termini per cell,

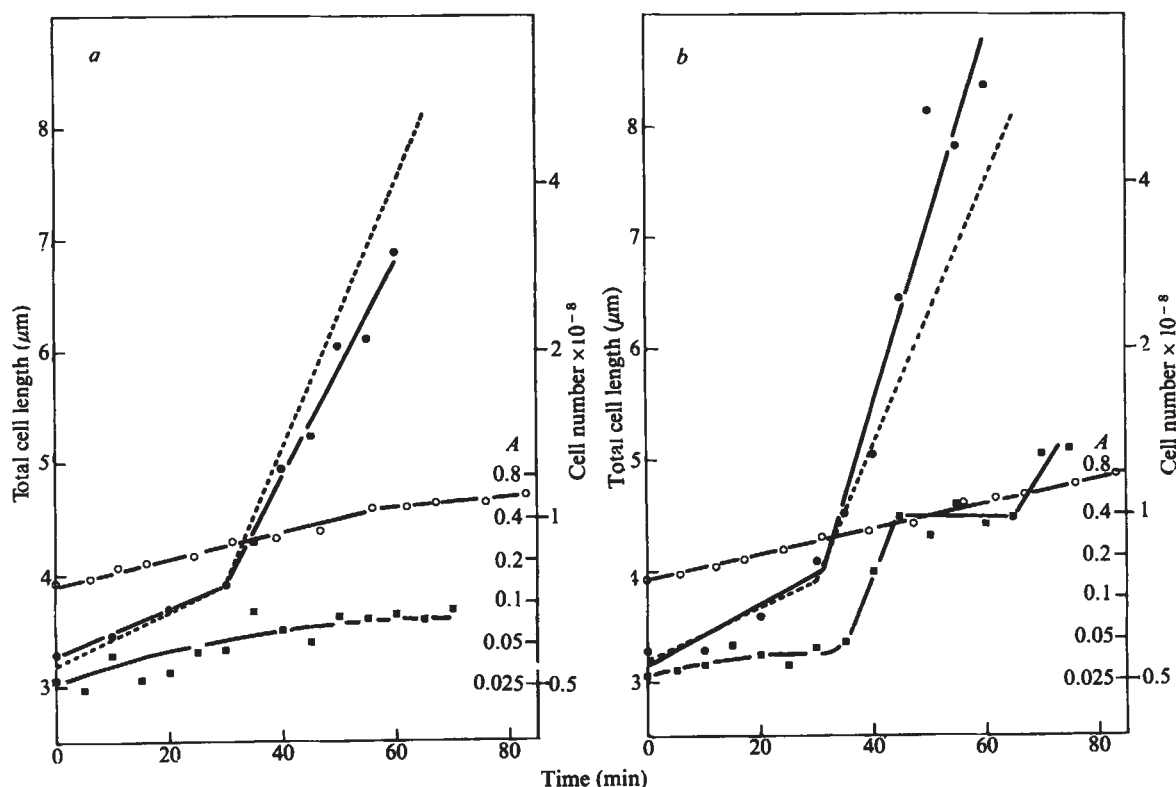


Fig. 4 Increase in total cell length, mass and number (symbols as in Fig. 3) following transfer of selected small cells to a richer medium, *a*, with or *b*, without further chromosome replication. A homogeneous fraction of small cells was selected by sucrose gradient centrifugation of a population of cells of *E. coli* B/r/1 T⁻, growing with a generation time of 56 min in medium 2 + thymine (50 μg ml⁻¹) and reinoculated into medium 7 (Table 1) *b*, with or *a*, without added thymine. In both cases the mass doubling time changed initially to 33 min but growth gradually slowed after about 50 min in the absence of thymine. DNA synthesis in this strain stops immediately after removal of thymine from the medium²⁵. In the presence of thymine, but not in its absence, synchronous cell division took place between 35 and 45 min and again after 65 min. (A small increase in cell number took place initially in both cultures, due to the separation of a proportion of cell pairs present in the selected fraction.) In both cultures total cell length increased initially at about the same rate and then increased abruptly at about 30 min. (This rate change was slightly higher than predicted in *b* and slightly lower than predicted in *a*, probably due to the general slowing down in growth after prolonged thymine deprivation, but the change in rate is clearly seen in both cases.)

caused by premature termination in the absence of cell growth, also does not affect the rate of cell elongation when growth is resumed. This experiment is described in Table 2.

Model of cell growth and division

We may summarise our observations as follows: (1) Cells divide about 20 min after reaching a particular length, 2λ , (which is about 2.8 μm for a K12 strain and for one B/r strain), independent of the growth rate of the cells. (2) The rate of cell elongation changes at about this same length. If growth conditions remain constant, the rate doubles at this point, whereas if growth conditions have been changed during the course of the preceding cycle, the growth rate changes at this same point to a value proportional to the cell's new rate of mass increase. (3) The critical length, 2λ , is twice the minimum theoretical length for cells. (4) Termination of chromosome replication, although necessary for subsequent cell division^{17,18}, does not control the change in rate of cell elongation.

A model consistent with these findings is that there is a minimum unit cell, of length λ μm , which has a fixed number of sites of elongation (perhaps one). Growth in length of such a unit cell takes place at these sites, at a rate proportional to the number of sites, as well as to the growth rate in cell mass, until a new unit length has been completed (at a total length of 2λ μm). At this point two things happen: septation is initiated between the two completed unit cells, and each of the two units begins to grow independently from new sets of growth sites.

The septation process, once initiated on the completion of the new unit, takes about 20 min to complete, independent of the growth rate. As a result, even though cells always initiate septation at a fixed length, the final length at the time of separation (and hence the average length of the population of cells) will be proportional to the growth rate. Thus, one of the oldest and simplest ideas about the control of cell division, namely that cells always begin to divide whenever they reach a certain critical size is, with minor modifications, seen to be consistent with our observations.

The rate at which each new unit cell grows would seem to be set at the time of its initiation. This rate is proportional to the growth rate in mass of the cell at the time it reaches length

2λ (which is the time when the new sets of growth sites are supposed to be activated). Because the rate of mass increase responds immediately to changes in growth conditions, no matter at which stage of the cycle the change is made¹⁸, the volume of cells changes with growth rate in a way which is different from that which we describe here for cell length. Consequently, cell diameter increases with growth rate such as to accommodate the larger increase in volume than length^{6,7}. The different ways in which growth in mass and in length respond to changes in growth conditions provide an explanation for the observed proportions of cells at different growth rates and also results in the change from one set of proportions to the next being made within a single cell cycle of a change in conditions (see Figs 3 and 4).

Sargent⁸ has recently reported that for several theoretical models of cell growth the length of cells of *Bacillus subtilis* changes in a way which is consistent with a doubling in rate of elongation at a constant cell length that is independent of growth rate. From his data, together with some measurements of our own, we find that this length is also close to twice that calculated for "minimum cells" (zero growth rate) and, interestingly, that this value is very close to that estimated for *E. coli*. The large size, relative to *E. coli*, of *B. subtilis* cells in most growth media can therefore be ascribed entirely to the long duration of the septation process (estimated at about 1 h) in this organism. Sargent also reported that cells of *B. subtilis* change only in length, without change in diameter, at different growth rates⁸. This implies that, unlike *E. coli*, the rate of elongation per growth site is likely to change immediately after a shift in growth rate.

The model of cell growth presented here seems to be inconsistent with the observations on the variation in cell length and width as a function of thymine concentration in cultures of thymine-requiring auxotrophs of *E. coli*^{6,7}. As the thymine concentration is reduced, the replication time of the chromosome is increased, so that termination and cell division are delayed and thus average cell mass is increased. Our present model predicts that this increase in cell volume should be the result solely of an increase in cell length, but the reported measurements show that it is in fact largely due to an increase in cell width. This observation led to the suggestion that the rate of cell elongation is proportional to the number of chromosome termini per cell^{6,7}. Our results, however, show no such relationship. We do not know the reason for this apparent discrepancy.

In our model, septation and the activation of new growth sites take place at a fixed cell length, 2λ . It is not implied however, that the trigger for these events is cell length *per se*. It may well be some correlated event, such as the completion of the required period of division protein synthesis¹⁸, which we have already suggested as the trigger for the initiation of the final septation process²⁰. But it is not possible to distinguish experimentally between the completion of a required period of protein synthesis and the attainment of the critical length as being the triggering event, because both are achieved at the same time and presumably involve similar biosynthetic reactions. The only correlated event which we have been able to eliminate is chromosome termination, because inhibition of DNA synthesis can be carried out specifically without immediately affecting cell growth. For the moment therefore, we can only point out the correlation and hope that further experiments will help to discover the proximate signal for the initiation of division.

We thank Millicent Masters and John F. Collins for advice.

Addendum: After this manuscript was submitted it came to our notice that N. B. Grover *et al.* have also measured cell length as a function of growth rate in *E. coli* with results which are indistinguishable from ours. They also conclude that their data agree with a model in which the rate of cell elongation doubles at about the time of chromosome termination.

Received May 24; accepted October 14, 1976.

¹ Marr, A. G., Harvey, R. J., and Trentini, W. C., *J. Bact.*, **91**, 2388-9 (1966).
² Schaechter, M., Maaloe, O., and Kjeldgaard, N. O., *J. gen. Microbiol.*, **19**, 592-606 (1958).

Table 2 Rate of elongation of "preterminated" cells and control cells

Time (min)	Calculated ratio (termination theory)	Observed ratios	
		$\bar{L}_p/\bar{L}_{c1}$	$\bar{L}_p/\bar{L}_{c2}$
0	1.00	1.01	1.01
40	1.29	0.97	0.95
60	1.38	—	0.97
80	1.30	0.90	0.96

Ratios of mean cell lengths in synchronous cultures of *E. coli* K12 "M6K", a *leu⁻thyA⁻(dra⁻ or drn⁻)* derivative of MB93 A81 (ref. 23) growing at 30 °C in medium 3 (Table 1). Small cells were selected by sucrose gradient centrifugation²⁴ and either (1) reinoculated into the same medium (control 1) or (2) starved for leucine for 120 min to allow termination of chromosome replication^{15,20} before readdition of leucine and resumption of growth (preterminated cells) or (3) starved for both leucine and thymine for 120 min to prevent termination before readdition of both supplements and resumption of growth (control 2). Synchronous division took place at about 80 min after reinoculation into fully supplemented medium in all three samples. Mean cell lengths were measured at the indicated times after resumption of cell growth ($\bar{L}_p$ for preterminated cells, $\bar{L}_{c1}$ for control 1; $\bar{L}_{c2}$ for control 2). The ratios of these lengths did not change significantly during the course of the experiments. The "calculated ratios" were based on the assumption that the rate of cell elongation would double at termination of chromosome replication. Termination in the control cells was assumed to take place 20 min before division. The maximum ratio of mean cell lengths would therefore be expected at 60 min. DNA synthesis was found to behave as expected, doubling once per cell cycle at the same time in both controls (following readdition of supplements), whereas it continued at the control rate for about 40 min in the leucine-starved cells and then stopped quickly. DNA synthesis in this sample, as expected, did not recommence immediately on readdition of leucine (although growth resumed immediately) but recommenced abruptly after about 20 more min.

- ³ Donachie, W. D., *Nature*, 219, 1077-9 (1968).
⁴ Helmstetter, C. E., Cooper, S., Pierucci, O., and Revelas, F., *Cold Spring Harb. Symp. quant. Biol.*, 23, 809-822 (1968).
⁵ Pritchard, R. H., Barth, P. T., and Collins, J., *Symp. Soc. gen. Microbiol.*, 19, 263-297 (1969).
⁶ Zaritsky, A., and Pritchard, R. H., *J. Bact.*, 114, 824-837 (1973).
⁷ Zaritsky, A., *J. theor. Biol.*, 54, 243-248 (1975).
⁸ Sargent, M. G., *J. Bact.*, 123, 7-19 (1975).
⁹ Donachie, W. D., and Begg, K. J., *Nature*, 227, 1220-1224 (1970).
¹⁰ Powell, E. O., *J. gen. Microbiol.*, 15, 492-511 (1956).
¹¹ Teather, R. M., Collins, J. F., and Donachie, W. D., *J. Bact.*, 118, 407-413 (1974).
¹² Woldringh, C. L., *J. Bact.*, 125, 248-257 (1976).
¹³ Schaechter, M., Williamson, J. P., Hood, J. R., and Koch, A. L., *J. gen. Microbiol.*, 29, 421-434 (1962).
¹⁴ Donachie, W. D., Jones, N. C., and Teather, R. M., *Symp. Soc. gen. Microbiol.*, 23, 9-44 (1973).
¹⁵ Maaaløe, O., and Kjeldgaard, N. O., *Control of Macromolecular Synthesis* (Benjamin, New York, 1966).
¹⁶ Pierucci, O., and Helmstetter, C. E., *Fedn Proc.*, 28, 1755-60 (1969).
¹⁷ Helmstetter, C. E., and Pierucci, O., *J. Bact.*, 95, 1627-1633 (1968).
¹⁸ Clark, D. J., *J. Bact.*, 96, 1214-1224 (1968).
¹⁹ Cooper, S., and Helmstetter, C. E., *J. molec. Biol.*, 31, 519-540 (1968).
²⁰ Jones, N. C., and Donachie, W. D., *Nature new Biol.*, 243, 100-103 (1973).
²¹ Hoffman, B., Messer, W., Schwarz, U., *J. supramolec. Struct.*, 1, 29-37 (1972).
²² Burdett, I. D. J., and Murray, R. G. E., *J. Bact.*, 119, 1039-1056 (1974).
²³ Smith, J. D., Barnett, L., Brenner, S., and Russell, R. L., *J. molec. Biol.*, 54, 1-14 (1970).
²⁴ Mitchison, J. M., and Vincent, W. S., *Nature*, 205, 987-9 (1965).
²⁵ Donachie, W. D., *J. Bact.*, 100, 260-8 (1969).

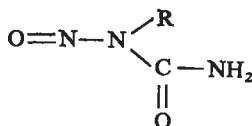
All oxygens in nucleic acids react with carcinogenic ethylating agents

B. Singer

Department of Molecular Biology and Virus Laboratory, and Space Sciences Laboratory, University of California, Berkeley, California 94720

Over 80% of ethylnitrosourea and ethylnitrosoguanidine modification of nucleic acids is on oxygens. The reactivity of oxygens (other than ribose and phosphate) in single-stranded RNA is: O^2 of C > O^2 of U > O^6 of G > O^4 of U. In double-stranded DNA the order is: O^2 of T = O^6 of G > O^4 of T > O^2 of C. Oxygen reactivity of single-stranded DNA resembles RNA. The glycosidic bond of O^2 -alkylpyrimidines is labilised.

ALKYLATING agents of the general structure



are potent carcinogens which do not require metabolic activation. The most studied of this type of nitroso compound are *N*-methyl-*N*-nitrosourea (MeNU) and *N*-ethyl-*N*-nitrosourea (EtNU). Another group of carcinogens which are also highly mutagenic are also nitroso derivatives, the best known of which are *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and *N*-ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG).

In recent years it has been found that the *N*-nitroso alkylating agents have an affinity for alkylating the oxygens of nucleic acids, in contrast to the "typical" alkylating agent (for example, dimethylsulphate) which is of low carcinogenicity or mutagenicity and which primarily alkylates ring nitrogens (reviewed by Singer, ref. 1). Following Loveless's suggestion that the O^6 of guanine could be alkylated in nucleic acids², a number of investigators did find that a significant proportion of the total alkylation was O^6 -alkyl G when nucleic acids were treated *in vitro* or *in vivo* with MeNU, MNNG and EtNU³⁻¹¹.

Due to the known lability of the alkyl group of O^6 -alkylguanosine, this derivative, and generally no other, was determined in enzyme digests of RNA rather than after hydrolysis with strong acid or base as customary for *N*-substituted derivatives. In DNA, either enzyme digestion or mild acid depurination was the method generally used to isolate that nucleoside or base without concomitant dealkylation. Singer and co-workers then found that enzyme digests of nucleic acids treated with MeNU or EtNU contained alkyl esters which were derived from alkylation of the internucleotide phosphodiester, and that alkyl phosphotriesters represented about two-thirds of the total bound alkyl groups in DNA and

RNA^{10,11}. Further examination of the alkyl products after enzyme digestion at neutral pH revealed that 10-15% of the alkylation of RNA by EtNU was on the ribose¹².

I now report that the newly described derivative, O^2 -ethylcytidine (O^2 -EtCyd)¹³, is a major product of neutral aqueous reaction of the single-stranded nucleic acids, TMV RNA and M13 DNA, with EtNU and ENNG. In double-stranded salmon sperm DNA this reaction is greatly suppressed and O^2 -ethylation of cytosine is only barely detectable. A further new finding is that both oxygens on uracil and thymine react with EtNU and ENNG. O^2 substitution of U in RNA and T in double-stranded DNA is greater than O^4 substitution, but in single-stranded DNA the O^2 and O^4 of T are equally reactive.

Techniques and their pitfalls

There are at least three reasons why *O*-alkylation of pyrimidines in nucleic acids has been overlooked until now. The first, and probably most important, is that one finds only what one looks for. Almost all of the *O*-alkyl pyrimidines have been prepared for the first time in this laboratory in the last year, and without authentic markers it is difficult to find and identify new products. Second, all *O*-alkyl pyrimidines are labile in acid and alkali so that losses (up to 100%) occur during the degradation of a nucleic acid unless neutral enzymatic digestion is used. Furthermore, separation of products in a digest must also be done under conditions which preserve the *O*-alkyl linkage. Third, mere coincidence of radioactivity with an ultraviolet absorbing marker in any system does not mean that the radioactive substance is completely or even partly identical with the marker.

The techniques used in this work which enabled us to avoid these pitfalls and to separate and quantitate the new derivatives discussed in this paper are all given in detail in preceding papers. Briefly they are: (1) the use of snake venom phosphodiesterase and phosphatases at pH 7.2 to digest alkylated polynucleotides and nucleic acids to nucleosides^{10-12,14,18}; (2) one-dimensional paper chromatography using butanol-ethanol-H₂O (80 : 10 : 25) (Solvent 1) to separate groups of possible products¹³⁻¹⁸; (3) additional separation techniques, for each group of possible products eluted from chromatograms, utilising the unique characteristics of the individual alkyl derivatives. For example, electrophoresis to separate products with different pK_a s^{15,16} or chromatography in borate-containing solvents to separate *O*-alkylnucleosides^{12,14}; (4) preparation and use as markers of all *N*- and *O*-monoethylated nucleosides¹⁰⁻¹⁸, and determination of their stability in conditions used for enzyme digestion, the R_f values in several neutral chromatographic systems, electrophoretic behaviour, rate of

acidic dealkylation, and so on; and (5) awareness of the fact that alkylation of pyrimidines in nucleic acids can diminish the stability of the glycosidic linkage.

Proof of formation of *O*²-ethylcytidine and *O*²-ethyldeoxycytidine

When alkylated single-stranded RNA, and poly(C) were enzyme-digested to nucleosides and chromatographed in a

neutral solvent, peaks of radioactivity co-chromatographed with authentic *O*²-ethylcytidine (Fig. 1) (data not shown for poly(C)). In the case of TMV RNA treated with EtNU or ENNG about 10% of the total radioactivity was in this area (A) (as shown in Fig. 1c and d), while with diethylsulphate (Et₂SO₄) reaction, no peak was exactly coincident with the marker (Fig. 1f) but about 2% of the radioactivity could be assigned to the marker area.

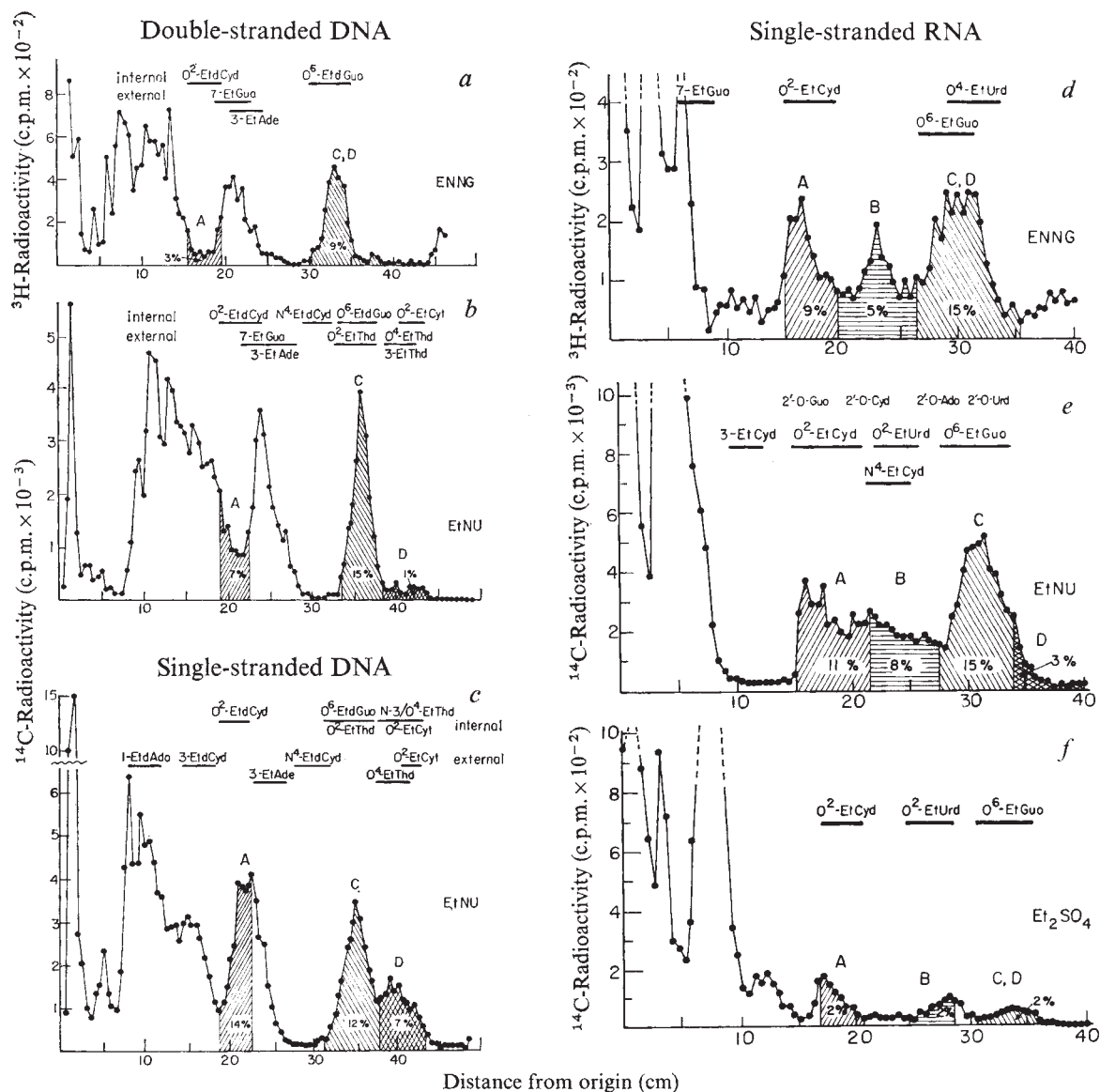


Fig. 1 Radioactivity profiles of chromatograms of enzyme digests of ethylated nucleic acids. 0.5–1 mg TMV RNA, salmon sperm DNA and M13 DNA were reacted in pH 6.1 or pH 7.3 cacodylate buffer 37 °C, 1–3 h, with ¹⁴C-ethylnitrosourea (Farbwerke Hoechst AG, 4.1 Ci mol⁻¹) or ³H-ethylnitrosoguanidine (Amersham/Searle, 110 Ci mol⁻¹) or ¹⁴C-diethylsulphate (ICN, 4.09 Ci mol⁻¹). The methods are described in detail in earlier papers^{10,11}. After precipitating with ethanol until constant specific activity was obtained, the specific activity of ethylated nucleic acids varied from 20–400 c.p.m. μg⁻¹. The lowest extent of reaction was consistently obtained with ethylnitrosoguanidine. The percent alkylation with each reagent was very similar for RNA and DNA^{10,11}. Ethylated RNA (50–200 μg) was digested with snake venom phosphodiesterase, acid and bacterial alkaline phosphatases using method I described by Kusmierek and Singer¹¹. 100–750 μg ethylated DNA in 0.1 M pH 7.2 Tris buffer containing 0.01 M MgCl₂ was digested with an equal weight of deoxyribonuclease, 37 °C, 4 h, then further digested with equal weights of snake venom phosphodiesterase, acid and bacterial alkaline phosphatases, 37 °C, 18 h. Singer and Kusmierek¹² have reported that these neutral digestion conditions permit recovery of labile *O*-alkyl nucleosides and other alkylated nucleosides or bases such as 7-alkyl G and 1-alkyl A which are ring-opened and rearranged, respectively, at higher pHs. After adding ultraviolet-absorbing markers (prepared in this laboratory^{10–18,20}), digests of alkylated RNA or DNA were applied to Whatman 3MM and developed 16–20 h in butanol–ethanol–water (80:10:25). The positions of the marker nucleosides or bases were noted, and the entire paper was cut into 0.5-cm strips (to which were added 5 ml toluene containing 14.3 g Omnifluor/3 kg toluene) for radioactivity counting. Panels a–f are typical profiles and in each panel the type of nucleic acid and reagent are given. The bars above the chromatograms represent the ultraviolet-absorbing areas of each marker and in general the internal and external markers have the same *R_f*. In e the approximate positions of 2'-*O*-ethyl nucleosides are also shown, based on their known *R_f* values¹². *O*²-EtCyd or *O*²-EtdCyd are termed 'A'; the area coinciding with *O*²-EtUrd, 'B'; the area coinciding with *O*⁶-EtdGuo, 'C'; and the area coinciding with *O*⁴-EtThd and *O*²-EtCyt or *O*⁴-EtUrd, 'D'. The percent of total radioactivity in each area is shown in the figures. After counting, paper strips from these and other areas were washed several times in toluene, dried and eluted with either water or methanol, then treated in various ways in order to separate the possible derivatives in each area (see text and Figs 2 and 4–6).

O^2 -EtCyd does not separate clearly in Solvent 1 from 2'- O -EtGuo or 1-EtGuo, and peak A could contain these two derivatives. However, both Guo derivatives have a pK_a of ~ 2 (ref. 19) while O^2 -EtCyd has a $pK_a > 9.5^{13}$. Consequently electrophoresis at $pH \sim 9$ was found to separate O^2 -EtCyd not only from derivatives of "low pK_a " but also from all other ethyl derivatives including those with a "high pK_a " such as 1-EtAdo and 3-EtCyd ($pK_a \sim 8.5$)¹⁹. Figure 2a illustrates that most of the radioactivity (from peaks *dA* and *eA*) coincides with O^2 -EtCyd upon electrophoresis in 0.1 M ammonium carbonate. The small peak of low pK_a was shown to be 2'- O -EtGuo ($\sim 1.8\%$ of total alkylation) when re-chromatographed in a borate-containing solvent¹². Further proof that peak A was primarily O^2 -EtCyd was that the radioactivity from the high pK_a area of the electrophoresis (1) co-chromatographed with the marker and could be separated from 1-EtGuo, the derivative with the closest R_f , by long chromatography in solvent 1 (Fig. 2b), and (2) the ^{14}C -labelled ethyl group was hydrolysed in $NHCl$ at the same rate as the marker was dealkylated to cytidine.

The ethylation of the O^2 of dCyd in DNA was studied using double-stranded salmon sperm DNA and single-stranded M13 DNA. As seen in Fig. 1a and b, enzyme digests of ENNG and EtNU-treated salmon sperm DNA did not exhibit a peak of radioactivity co-chromatographing with O^2 -EtdCyd. The radioactivity, termed peak A, which really represented a minimum, was analysed using similar methods to those used to detect O^2 -EtCyd and it was found that O^2 -EtdCyd was less than 0.3% of the total radioactivity. The large peak moving slightly ahead of the marker O^2 -EtdCyd was identified by two-dimensional chromatography¹¹ as containing 7-EtGua and 3-EtAde (3:1); both bases were also in peak A and their presence illustrates the lability of the glycosidic bond of alkylpurine deoxynucleosides even at $pH 7.2$.

In studying the stability of the ethyl group in O^2 -EtdCyd (prepared from dCyd as described for the preparation of O^2 -EtCyd from Cyd¹³) it was found that heating in 0.1 N HCl at $70^\circ C$ for 30 min or at $pH 7.0$, $100^\circ C$ 1 h, caused complete depyrimidation without loss of the ethyl group and the product was O^2 -ethylcytosine (Fig. 3). These conditions are often

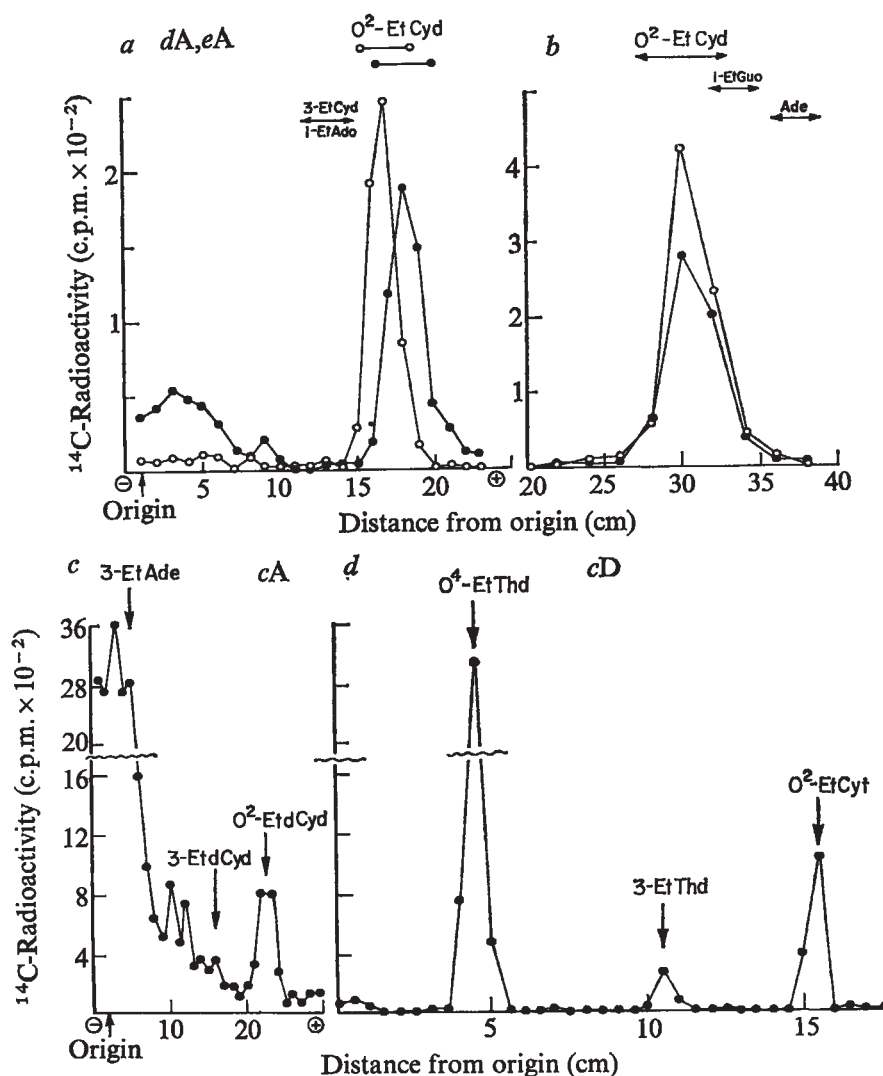


Fig. 2 Separation and identification of O^2 -EtCyd or O^2 -EtdCyd and O^2 -EtCyt. *a*, Superimposed radioactivity profiles of electrophoretograms of peak A from Fig. 1d and e (termed *dA*, *eA*). Material eluted from the O^2 -EtCyd area of EtNU- and ENNG-treated TMV RNA was separately electrophoresed in 0.1 M $(NH_4)_2CO_3$ ($pH \sim 9$)¹³. The electrophoretogram was cut in 101 cm strips and the radioactivity counted. The open circles are from *dA* and the closed circles from *eA*. The position of the internal O^2 -EtCyd for each electrophoretogram is similarly indicated. The mobility, at this pH , of O^2 -EtCyd was higher than any other ethylated nucleoside as shown by the separation of O^2 -EtCyd from 3-EtCyd ($pK_a 8.4$) and 1-EtAdo ($pK_a \sim 8.5$) *b*, Superimposed radioactivity profiles of a portion of chromatograms of the material co-electrophoresing with O^2 -EtCyd in *a*. After elution, samples were chromatographed in the same solvent as in Fig. 1 except that the chromatogram was developed 48 h. O^2 -EtCyd was internal while 1-EtGuo and Ade were used as external markers since with 16–20 h chromatography these derivatives do not separate from O^2 -EtCyd (data not shown) while with longer development their movement differs. *c*, Radioactivity profile of an electrophoretogram of peak A from Fig. 1c (termed *cA*). Material eluted from the O^2 -EtdCyd area of EtNU-treated M13 DNA was electrophoresed in 0.1 M $(NH_4)_2CO_3$ ($pH \sim 9$). O^2 -EtdCyd is completely separated from 3-EtAde ($pK_a 6.5$) and 3-EtdCyd ($pK_a 8.5$) as well as all derivatives with a $pK_a < 6.5$. The radioactivity co-electrophoresing with O^2 -EtdCyd had the same acid stability as the marker. *d*, Radioactivity profile of a silica gel thin-layer chromatogram of peak D from Fig. 1c (termed *cD*). This peak from EtNU-treated M13 DNA which contained ultraviolet-absorbing markers of O^4 -EtThd, 3-EtThd and O^2 -EtCyt was developed in acetone–benzene (2:1), then cut in 0.5 cm strips and the radioactivity counted.

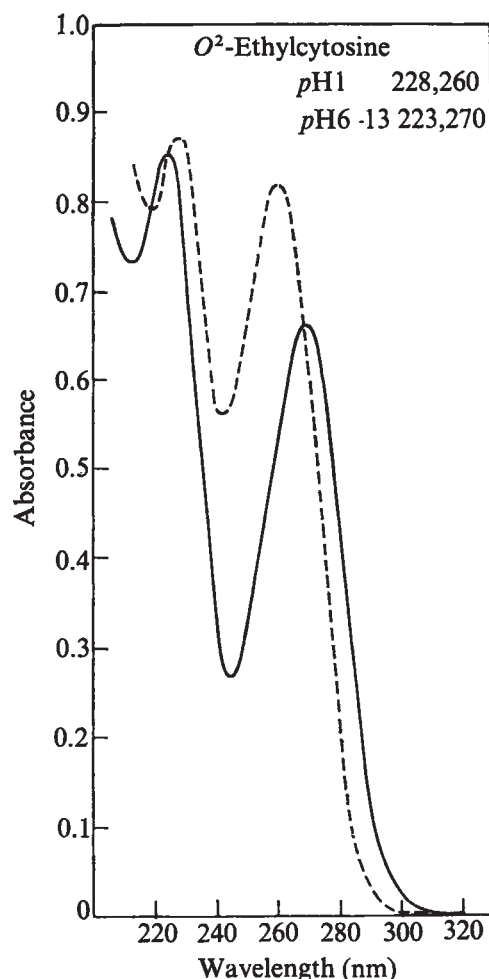


Fig. 3 Ultraviolet absorption spectra of O^2 -ethylcytosine at pH 1 (---) and pH 6-13 (—). The absorption maxima are given in the figure. O^2 -Ethylcytosine was prepared by the following method. To 30 mg cytosine in 10 ml methanol was added a solution of diazoethane in ether prepared according to the procedure of Robins *et al.*³³ for the preparation of diazo-methane (solution A). Several additions of several millilitres each were needed to saturate the solution. The ether was evaporated off and the remaining solution chromatographed on Whatnam 3MM in butanol-ethanol-water (80:10:25) to separate the several derivatives. O^2 -Ethylcytosine has in this system an R_f 0.88 and moves ahead of all other ultraviolet absorbing compounds. The yield from 30 mg cytosine is about 300 absorbance units. O^2 -methylcytosine was prepared with a lesser yield using diazomethane. The methyl derivative, obtained by a different method, has been described by Sukhorukov *et al.*³⁴, although the second λ_{max} (228 in acid and 223 in alkali) was not shown by them probably due to impurities in their sample. The reported pK_a of O^2 -methylcytosine is 5.41 (ref. 36).

used for depurinating alkylated DNA and would be likely to depyrimidate O^2 -EtdCyd. Since a substantial proportion of 7-EtdGuo and 3-EtdAdo was depurinated during pH 7.2 enzyme digestion, it seemed reasonable to suspect that, in addition to 7-EtdGua and 3-EtdAde, O^2 -EtCyt would also be present and by R_f would be in peak D. This was shown by electrophoresis of peak D (Fig. 1c) at pH 5.5, where O^2 -alkylcytosine which has a reported pK_a of 5.4 (ref. 19) migrates while the other possible products, O^4 -EtThd and 3-EtThd, do not move at this pH. Although O^2 -EtCyt could be identified, the amount was extremely low.

Identification of O^2 -EtdCyd and O^2 -EtCyt was unequivocal in EtNU-treated single-stranded M13 DNA. First, although peak A (Fig. 1c) was not completely coincident with O^2 -EtdCyd, the radioactivity sharply increased in the area of the marker,

rather than decreasing as was the case for double-stranded DNA. Electrophoresis at pH ~9 of the shaded area in peak A clearly resolved O^2 -EtdCyd from other possible neighbouring ethyl nucleosides of pK_a 6.5 and pK_a ~8.5 (Fig. 2c), and there was a distinct peak of radioactivity co-electrophoresing with the authentic nucleoside. On the basis of earlier experiments it was expected that O^2 -EtCyt would also be formed as a result of deprimidation during enzyme digestion. Peak D contained as internal markers O^2 -EtThd, 3-EtThd and O^2 -EtCyd, which were resolved by silica gel thin layer chromatography. A distinct and sharp peak of radioactivity coincided with O^2 -EtCyt (Fig. 2d) and other similarly sharp peaks were found with the other two markers.

Proof of formation of O^2 - and O^4 -ethyluridine and O^2 - and O^4 -ethylthymidine

The R_f of O^2 -EtUrd in the chromatographic system used in Fig. 1 is very similar to the R_f s of 2'- O -EtCyd and N^4 -EtCyd and the peak from RNA digests designed as 'B' in Fig. 1d-f could contain all three derivatives. Similarly, O^4 -EtUrd is coincident with 3-EtUrd (peak D) and not completely separable from O^6 -EtGuo and 2'- O -EtUrd (peak C) (Fig. 1d-f). Several techniques were used to resolve the derivatives in peaks B, C and D; the most useful of these were electrophoresis in pH 3.0 sodium citrate to separate O^2 -EtUrd (pK_a < 1) from N^4 -EtCyd and 2'- O -EtCyd (pK_a ~4.2)¹⁹ and thin-layer chromatography²⁰ which can separate O^4 -EtUrd from O^6 -EtGuo, 3-EtUrd and 2'- O -EtUrd in peaks C and D. Figures 4b and 5b illustrate these separations. The areas of radioactivity co-chromatographing with O^4 -EtUrd (shaded area of Fig. 4b) and co-electrophoresing with O^2 -EtUrd (shaded area of Fig. 4a) were each eluted and in both cases all radioactivity was volatilised after treatment with N HCl (100 °C, 60 min) and the ultraviolet-absorbing markers were converted to uridine, thus confirming that the original derivatives were oxygen substituted.

O^2 -EtThd has the same R_f as O^6 -EtdGuo (peak C) (Fig. 1a-c) and the two nucleosides also behave similarly in that both lose their ethyl group upon hydrolysis in N HCl at 100 °C, 60 min. However, thin-layer chromatography²⁰ clearly separates them (Figs 5a and 6a) as well as O^4 -EtThd and 3-EtThd which are mainly in peak D (Fig. 6b). Thin-layer chromatography of

Table 1 Proportion of ethyl products from hydrolysates of nucleic acids reacted with ^{14}C -ethylnitrosourea

Derivative	TMV RNA	M13 DNA	Salmon sperm DNA
	Total ethylation* (%)		
O^6 -EtGuo or O^6 -EtGuo	5†	7	7‡
O^2 -EtUrd	7		
O^2 -EtThd		5	6
O^4 -EtUrd	3		
O^4 -EtThd		5	1.4
O^2 -EtCyd	8		
O^2 -EtdCyd + O^2 -EtCyt		3	0.5
Total	23%	20%	15%
2'- O -EtAdo, -Guo, -Cyd, -Urd	12 ± 2§	—	—
Ethylphosphotriester	65¶		70
N-Ethyl	16		17

*See Fig. 1 for methods. The per cent ethylation in different experiments varied from 0.5 to 1.5%.

†Singer and Fraenkel-Conrat¹⁰ reported 12% O^6 -EtGuo on the basis of coincidence of radioactivity with the authentic marker and volatility of most of the ethyl groups in N HCl 100% 1 h. The present data show that this peak also contains O^4 -EtUrd and 2'- O -EtUrd.

‡Sun and Singer¹¹ reported 12% O^6 -EtGuo on the same basis as footnote †. The present data, if O^2 -EtThd and O^6 -EtdGuo are added, give a similar value for this peak.

§Data from Singer and Kusmierek¹².

¶Data from Singer and Fraenkel-Conrat¹⁰. The amount of esterification was determined by distillation of ethanol after N HCl hydrolysis. Ethanol calculated to be derived from O^6 -EtGuo was subtracted. If now the volatile ethyl groups on the uridine ring are subtracted as well, the calculated amount of riboalkylphosphotriester is about 54%.

||Data from Sun and Singer¹¹.

combined peaks C and D further illustrates these separations (Fig. 5a). One additional proof of the identity of O^2 -EtThd is the great instability of both the alkyl and the glycosidic bond²⁰; thus hydrolysis in 0.1 N HCl at 70 °C for 30 min almost quantitatively converts this derivative to thymine, while O^6 -EtdGuo is only depurinated, becoming O^6 -EtGua.

O^4 -EtThd and 3-EtThd (peak D) do not separate from O^2 -EtCyd in solvent 1 although the R_f s are somewhat different (Fig. 1b and c). They all separate completely upon thin-layer chromatography, with O^2 -EtCyt moving ahead of the other derivatives (Fig. 2d). In addition, O^2 -EtCyt ($pK_a \sim 5.4$) can

basis of its symmetry and sharpness, contains two major and one minor component (Fig. 6a).

There are two points to be noted before quantitative conclusions are drawn from the data on the figures and Table 1. First, although the alkylated nucleic acids were digested for long periods of time with large amounts of enzymes, from 10 to 40% of the total radioactivity remained at the origin. This material, which was not due to generally incomplete digestion since there was no associated ultraviolet absorbancy, did not elute from the paper even with 0.1 N HCl or long incubation at 37 °C. Since the origin material is labelled, it obviously

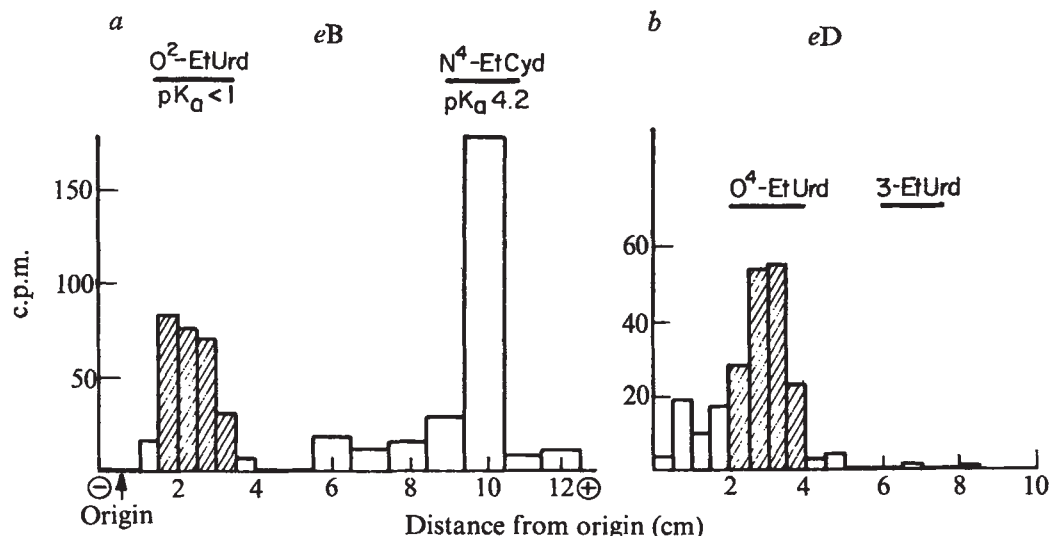


Fig. 4 Separation and identification of O^2 -EtUrd and O^4 -EtUrd. *a*, Radioactivity profile of an electrophoretogram of peak B from Fig. 1e (termed eB). Material eluted from the O^2 -EtUrd area of EtNU-treated TMV RNA was electrophoresed in 0.1 M pH 3 sodium citrate buffer. O^2 -EtUrd does not migrate at this pH while nucleosides which also could be in peak B, for example, N^4 -EtCyd, move very far. *b*, Radioactivity profile of a silica gel thin-layer chromatogram of peak D from Fig. 1e (termed eD). Material eluted from the O^4 -EtUrd area of EtNU-treated TMV RNA, which also contained a marker of 3-EtUrd, was chromatographed using benzene-acetone (2:1). See Fig. 5b for separation of O^4 -EtUrd from O^6 -EtGuo, 2'- O -EtUrd, and 3-EtUrd.

be separated from the thymidine derivatives by electrophoresis in pH 5.7 formate buffer.

Quantitative aspects of *O*-ethylation

In Fig. 1, the percentages given in shaded areas A–D refer to percentage of total recovered radioactivity. In the case of RNA treated with ENNG or EtNU, most of the radioactivity in these areas could be assigned to a known *O*-alkyl derivative. Thus peak A is essentially O^2 -EtCyd (Fig. 2), peak B is about half O^2 -EtUrd (Fig. 4a), peak C contains O^6 -EtGuo and 2'- O -EtUrd while peak D is O^4 -EtUrd (Fig. 4b and Fig. 5b). Diethylsulphate-treated RNA (Fig. 1f) was very low in all these derivatives and none of these peaks contained more than 0.5% (of the total radioactivity) of an identifiable derivative. Nevertheless, O^2 -EtCyd was identified in peak A and O^6 -EtGuo and O^4 -EtUrd in peaks C, D.

EtNU-treated single-stranded DNA resembled EtNU-treated single-stranded RNA in the chromatographic profile and about 10% of peak A (Fig. 1c) is O^2 -EtdCyd (Fig. 2c). Peak C is about 40% O^2 -EtThd, 55% O^6 -EtdGuo and the remaining 5% is O^4 -EtThd coming from peak D where it is a major component. Peak D comprises 70% O^4 -EtThd, 24% O^2 -EtCyt and 6% 3-EtThd.

Area A of EtNU- and ENNG-treated salmon sperm DNA (Fig. 1a and b) contains primarily N-ethyl products with a trace of O^2 -EtdCyd. Peak C is almost equally divided into O^2 -EtThd and O^6 -EtGuo (Fig. 6a) while peak D is mostly O^4 -EtThd (Fig. 6b). It also can be seen that neither O^4 -EtThd nor O^4 -EtUrd is totally separated from peak C. Another, and more important, fact to be noted, is that a peak, such as C in Fig. 1b, which might be surmised to be homogeneous on the

represents bound alkyl groups and cannot be disregarded. But at this time I have no further information. (One possibly pertinent point is that alkylated poly(U) and poly(C) are almost completely digested using the same enzymes as judged by the fact that less than 5% of the radioactivity is at the origin.) All percentages of known derivatives are based on total radioactivity on the paper and do include the origin. Consequently it is possible that any derivative may remain partly in an enzyme-resistant oligonucleotide at the origin.

Second, *O*-alkyl derivatives are relatively labile and particularly in the case of DNA the percentage of O^2 -EtThd and O^2 -EtdCyd should be taken as minima since these two derivatives are depyrimidated easily and may be lost during reaction.

Given these caveats, Table 1 presents data for EtNU-treated nucleic acids, averaged from several experiments, which show that O^2 -EtUrd or O^2 -EtThd is a major derivative in EtNU-treated TMV RNA, salmon sperm DNA and M13 DNA. O^4 -EtUrd or O^4 -EtThd is formed in lesser amounts but still exceeds the amount of most N-ethyl derivatives. O^2 -EtCyd is formed to a greater extent than O^6 -EtGuo in EtNU- or ENNG-treated TMV RNA. Formation of this derivative is suppressed in double-stranded DNA but when the DNA is single stranded, considerably more reaction takes place. Data for the alkylation products of ENNG-treated nucleic acids are very similar and are not given in tabular form.

A summary of the relative reactivity of various sites in single- and double-stranded nucleic acids treated with EtNU or ENNG is as follows:

Double-stranded DNA:

Phosphate > N7-G > O^2 -T, O^6 -G, > N3-A > O^4 -T > O^2 -C, other N-

Single-stranded RNA:

Phosphate > Ribose, N^7 -G > O^2 -C > O^2 -U, O^6 -G > O^4 -U > N^1 -A, N^3 -C > other N-

The two homo-polynucleotides which were treated with the same reagents are also modified primarily on the oxygens and the data are given below for comparison. The work on poly(U) is published separately¹⁴ but the data for poly(C) are from the present series of experiments.

Poly(U): Phosphate > O^4 > O^2 , Ribose $\gg$ N3

Poly(C): O^2 > N^3 , N^4 > Phosphate > Ribose

We have routinely in the recent past used N HCl volatility as a semi-quantitative test for *O*-alkylation in RNA and HClO₄ volatility for *O*-alkylation in DNA. The alkyl groups of all of the derivatives discussed in this paper are hydrolysed by the HCl treatment except for ribose alkyl nucleosides which require HClO₄. Therefore those alkylating agents which do not with acid yield an alcohol may rather safely be presumed not to modify oxygens significantly. As an example, Singer and Fraenkel-Conrat¹⁰ reported 1% ethanol upon HCl hydrolysis of Et₂SO₄-treated RNA and now Fig. 1f illustrates that little *O*-alkylation of pyrimidines occurs, thus serving to validate analytical data obtained with alkylsulphate and alkylalkanesulphonates before our knowledge of *O*-alkylation of pyrimidines^{10,11}.

Possible biological consequences of *O*-alkylation

The study of the alkylation of nucleic acids started from the observations that alkylating agents such as mustard gas were

inactivating to DNA containing viruses²¹ and that mustard gas was mutagenic²². The principal site of reaction with mustard gas and many other alkylating agents was found to be the N^7 of guanine¹ and from then on the presence of 7-alkylguanine was presumed to be the significant event in causing a biological effect. Since alkylation of the N^7 did not change the base-pairing properties of G^{23,24}, its effects in DNA were generally attributed to the rapid depurination of 7-alkyl G which further led to chain breakage by β -elimination.

As time went on the simple relationship between the amount of 7-alkylguanine and biological change eroded. This became particularly apparent when alkylation-induced carcinogenesis was studied and ethylating agents were more carcinogenic in rats than the analogous methylating agents while the amount of 7-ethyl G was a small fraction of the amount of 7-methyl G¹.

When the highly carcinogenic *N*-nitroso compounds were used to modify nucleic acids *in vitro* or *in vivo* another derivative was found, *O*⁶-alkylguanine, which represented a few per cent of the total alkyl groups³⁻¹¹. When reactions were performed with alkylating agents of low carcinogenicity, this derivative was not detected or was present in trace amounts^{3,4,6,10,11,25}. Thus the effects of the presence of *O*⁶-alkylguanine became of prime interest since there seemed to be a correlation between its formation and the carcinogenicity of the alkylating agent. In contrast to 7-alkyl G, *O*⁶-alkyl G in polymers has been shown to mispair²⁴ and thus could be considered mutagenic. There is also good evidence that there is an enzymatic excision of *O*⁶-alkyl G from DNA²⁶ and in rats given MeNU or EtNU under conditions where only brain tumours develop, *O*⁶-alkyl G is removed from the brain very much more slowly than from

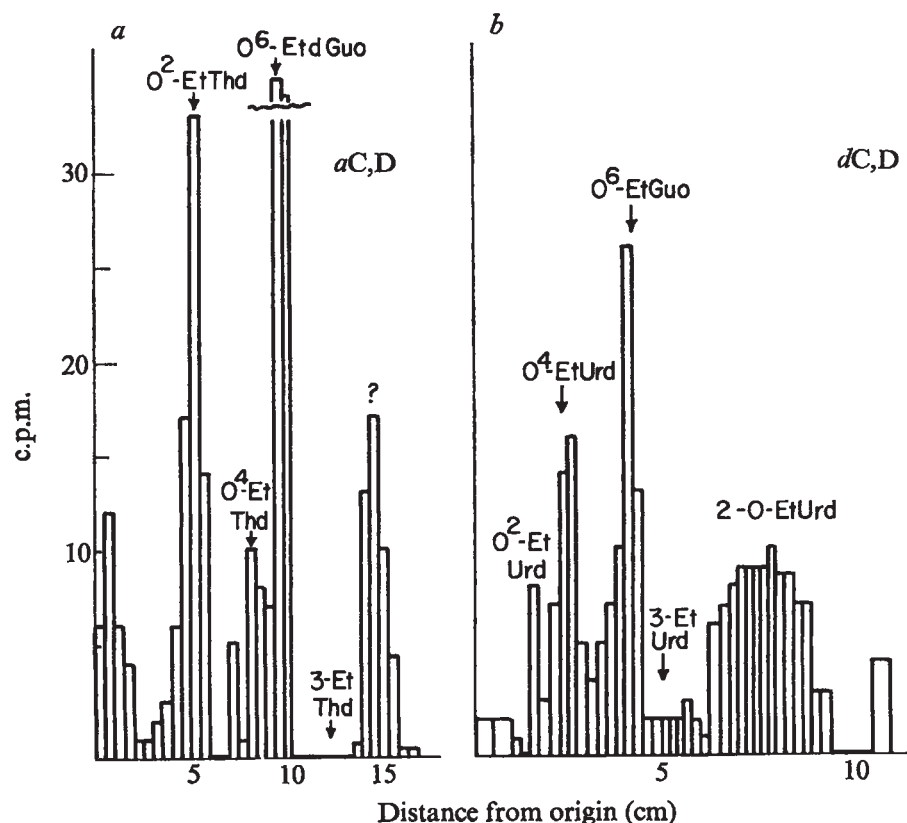


Fig. 5 Separation and identification of *O*⁶-EtdGuo, *O*²-EtThd and *O*⁴-EtThd from DNA and the corresponding ribo derivatives from RNA. Radioactivity profile of silica gel thin-layer chromatograms developed with benzene-acetone (2:1). The positions of internal ultra-violet absorbing derivatives are shown by arrows. *a*, Combined peaks C, D from Fig. 1a (termed aC,D). This material from ENNG-treated salmon sperm DNA contained internal markers of *O*²-EtThd, *O*⁴-EtThd, *O*⁶-EtdGuo and 3-EtThd. *b*, Combined peaks C, D from Fig. 1d (termed dC,D). This material from ENNG-treated TMV RNA contained internal markers of *O*⁴-EtUrd, *O*⁶-EtGuo, 3-EtUrd and 2'-*O*-EtUrd.

other tissues which do not develop neoplasms^{6,9,27-31}. It can be postulated that neoplasms develop in the brain because O^6 -alkyl G remains in the DNA and mispairing occurs.

The same reagents which alkylate the O^6 of G have been shown in this paper to react with all oxygens of the pyrimidines (Table 1). In quantitative terms the O^2 of T in double-stranded DNA reacts to the same extent as does the O^6 of G. Since the alkylation of the O^2 position of T greatly labilises the glycosidic linkage, it is likely that O^2 -alkyl T may depyrimidinate spontaneously or enzymatically leaving a gap which may or may not

or translation. Whether carcinogenesis can be explained solely in terms of mutation is doubtful since mutation is a frequent event and easily produced in cells while carcinogenesis is rare and not yet a test-tube phenomenon.

In conclusion, the *N*-nitroso alkylating agents are powerful carcinogens and alkylate all oxygens in nucleic acids. Other alkylating agents of lesser carcinogenicity do not have an affinity for oxygens. It is tempting to hypothesise that the carcinogenicity of an alkylating agent can be assessed by its reactivity towards oxygens, but I believe this is still premature.

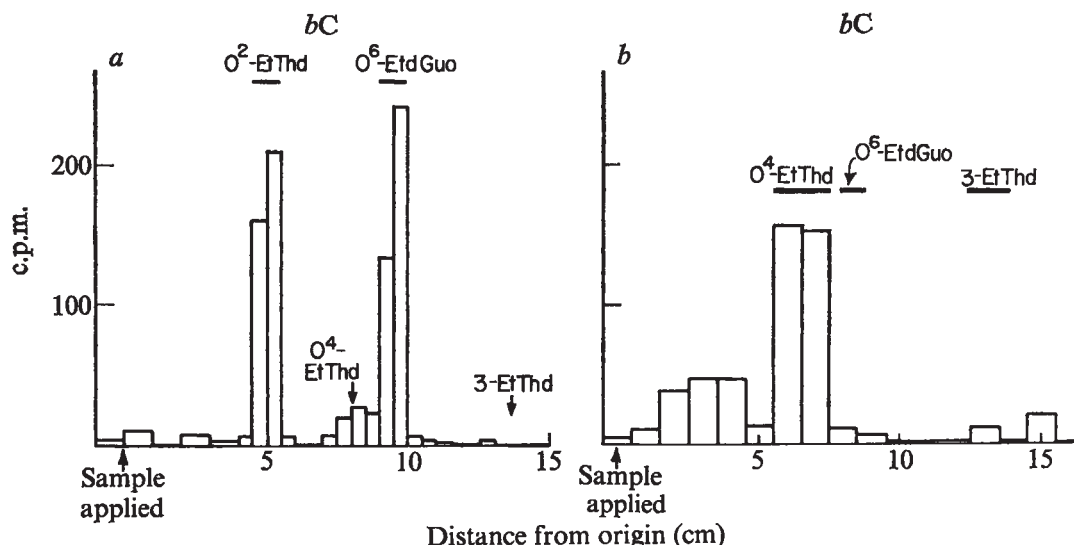


Fig. 6 Separation and identification of O^6 -EtdGuo, O^2 -EtThd and O^4 -EtThd. Radioactivity profiles of silica gel thin-layer chromatograms developed with benzene-acetone (2:1). Peaks C and D, from EtNU-treated salmon sperm DNA (termed 2C and 2D) were separately chromatographed. The position of the internal ultraviolet-absorbancy derivatives are shown by bars or arrows. a, Peak bC containing internal markers of O^2 -EtThd, O^4 -EtdGuo and 3-EtThd. See Fig. 5, left panel, for a similar chromatogram of combined peaks C and D.

be repaired. In O^2 -alkyl U or T there is no proton at the *N*-3 which means that normal base-pairing with A cannot occur.

Alkylation of the O^4 of U or T may also lead to mispairing since these derivatives also cannot pair normally to A but instead may tend to base pair with G and thus be considered pro-mutagens. While the amount of O^4 -EtThd is relatively low in double-stranded DNA, it is considerable in single-stranded nucleic acids. This is perhaps the place to emphasise that the study of alkylation of "double-stranded" DNA must also consider that replicating DNA contains single-stranded segments and this portion of the molecule will react in a similar manner to any single-stranded polynucleotide, be it RNA or DNA.

The O^2 of C in double-stranded DNA is quite unreactive but in single-stranded DNA the reactivity is greatly increased, and in RNA the amounts of O^2 -alkyl C exceeds that of any derivative modified on the base. This modification in DNA destabilises the glycosidic bond and even with the most gentle handling, half of the derivative is found as the base, O^2 -ethylcytosine. This readiness to depyrimidinate may be occurring continually in the cell with the effect of creating deletions. In RNA where the glycosidic linkage is stable, O^2 -alkylation of C forms a quaternary structure and it is predicted that there will be random mispairing of bases.

Two other significant oxygen alkylations by *N*-nitroso compounds should be mentioned although they are the subjects of separate papers. In RNA ribose is alkylated¹², a modification which may affect mRNA function, and in both DNA and RNA alkyl phosphotriesters are formed (50-70% of all alkylation)^{10,11} which we have shown to be lethal³². Thus all types of alkylation of a nucleic acid are likely to have biological consequences.

These are probably all related in some way to genome mutation or messenger alteration, that is, a change in transcription

I would like to acknowledge the contributions of my collaborators, Drs L. Sun and J. T. Kuśmierk, in much of this work. Dr H. Fraenkel-Conrat has been a constant source of advice and criticism.

Received June 25; accepted September 27, 1976.

- Singer, B., *Progr. Nucl. Acids Res. molec. Biol.*, **15**, 219-284, 330-332 (1975).
- Loveless, A., *Nature*, **223**, 206-207 (1969).
- Lawley, P. D., and Thatcher, C. J., *Biochem. J.*, **116**, 693-707 (1970).
- Lawley, P. D., and Shah, S. A., *Biochem. J.*, **128**, 117-132 (1972).
- Lawley, P. D., Orr, D. J., Shah, S. A., Farmer, P. B., and Jarman, M., *Biochem. J.*, **135**, 193-201 (1973).
- Walker, I. G., and Ewart, D. F., *Mutat. Res.*, **19**, 331-341 (1973).
- Kleihues, P., and Magee, P. N., *J. Neurochem.*, **20**, 595-600 (1973).
- Goth, R., and Rajewsky, M. F., *Z. Krebsforsch.*, **82**, 37-64 (1974).
- Margison, G. P., and Kleihues, P., *Biochem. J.*, **148**, 521-525 (1975).
- Singer, B., and Fraenkel-Conrat, H., *Biochemistry*, **14**, 772-782 (1975).
- Singer, B., and Singer, B., *Biochemistry*, **14**, 1795-1802 (1975).
- Singer, B., and Kuśmierk, J. T., *Biochemistry* (in the press).
- Singer, B., *FEBS Lett.*, **63**, 85-88 (1976).
- Kuśmierk, J. T., and Singer, B., *Biochim. biophys. Acta* (in the press).
- Singer, B., *Biochemistry*, **12**, 3939-3947 (1972).
- Sun, L., and Singer, B., *Biochemistry*, **13**, 1905-1913 (1974).
- Singer, B., Sun, L., and Fraenkel-Conrat, H., *Biochemistry*, **13**, 1913-1920 (1974).
- Singer, B., *Biochemistry*, **14**, 4353-4357 (1975).
- Singer, B., in *Handbook of Biochemistry and Molecular Biology* (edit. by Fasman, G. D.), *Nucleic Acids*, **1**, 409 (1975).
- Kuśmierk, J. T., and Singer, B., *Nucleic Acids Res.*, **3**, 989-1000 (1976).
- Herriott, R. M., *J. gen. Physiol.*, **32**, 221-239 (1948).
- Auerbach, C., and Robson, J. M., *Nature*, **157**, 302 (1946).
- Ludlum, D. B., *J. biol. Chem.*, **245**, 477-482 (1970).
- Gerchman, L. L., and Ludlum, D. B., *Biochim. biophys. Acta*, **308**, 310-316 (1973).
- Craddock, V. M., *Biochim. biophys. Acta*, **312**, 202-210 (1973).
- Kirtikar, D. M., and Goldthwait, D. A., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 2022-2026 (1974).
- Goth, R., and Rajewsky, M. F., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 639-643 (1974).
- Kleihues, P., and Margison, G. P., *J. natn. Cancer Inst.*, **53**, 1839-1841 (1974).
- Rajewsky, M. F., Goth, R., Laerum, O. D., Biessmann, H., and Hülser, D. F., *Vith Int. Symp. Princess Takamatsu Cancer Research Fund* (in the press).
- Nicoll, J. W., Swann, P. F., and Pegg, A. E., *Nature*, **254**, 261-262 (1975).
- Kleihues, P., and Margison, G. P., *Nature*, **259**, 153-155 (1976).
- Singer, B., Sun, L., and Fraenkel-Conrat, H., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 2232-2236 (1975).
- Robins, M. J., Naik, S. R., and Lee, A. S. K., *J. Org. Chem.*, **39**, 1891-1899 (1974).
- Sukhorukov, B. I., Gukovskaya, A. S., Sukhoruchkina, L. V., and Laurrenova, G. I., *Biophysica*, **17**, 5-11 (1972).

letters to nature

Direct test of the constancy of fundamental nuclear constants

THE possibility that fundamental nuclear constants may vary slowly while the Universe expands has been discussed by several authors¹⁻⁵. I try here to show that the well known resonance properties of the 'heavy nucleus plus slow neutron' system make it a sensitive 'receiver', sharply tuned to the current values of nuclear constants.

What are the restrictions, imposed by experiment that during the time interval ΔT the resonance energy shift had not exceeded ΔE_{exp} ? Simple estimates of residual interaction matrix elements suggest that for the overwhelming majority of compound nucleus resonances one should expect their shifts to be not less than the single-particle resonance shift ΔE_0 . The latter is connected with the relative change in strong coupling constant g_s ,

$$\frac{\Delta E_0}{V_0} \approx \frac{\Delta g_s}{g_s} \quad (1)$$

where V_0 denotes the depth of the nuclear potential well. Assuming $dg_s/dt = \text{constant}$, we get a restriction

$$\frac{1}{g_s} \left| \frac{dg_s}{dt} \right| \lesssim \frac{1}{\Delta T} \frac{\Delta E_{\text{exp}}}{V_0} \quad (2)$$

The positions of many low lying resonances have been known to an accuracy of 10^{-3} eV for quite a time⁶. Assuming $V_0 \approx 50$ Mev (ref. 7) and $\Delta T \approx 10$ yr we derive

$$\frac{1}{g_s} \left| \frac{dg_s}{dt} \right| \lesssim 2 \times 10^{-12} \text{ yr}^{-1} \quad (3)$$

which is the same result as that established by Davies on the basis of Dyson's cosmological argument⁵.

The Coulomb force increases the average internucleon distance by $\sim 2.5\%$ for $A \approx 150$ (ref. 7). Thus we obtain an estimate for the Coulomb coupling constant α 20 times higher than

Table 1 Comparison of upper bounds of the variation of nuclear constants

	Dyson, Davies	Present work
$1/g_s dg_s/dt (\text{yr}^{-1})$	2×10^{-12}	5×10^{-10}
$1/\alpha d\alpha/dt (\text{yr}^{-1})$	2×10^{-14}	10^{-17}
$1/g_w dg_w/dt (\text{yr}^{-1})$	10^{-10}	2×10^{-12}

from equation (2). The weak interaction contribution to the total energy of the nucleus is $\sim 10^{-5}(\mu/m)^2$, where μ and m are the pion and nucleon mass, respectively⁸. The upper bound on the time variation of the weak coupling constant g_w is therefore 5×10^6 times higher than for g_s .

The low lying resonance parameters determine the capture cross section for slow neutrons. So data on thermal cross section values in the remote past are of great interest. The recently discovered traces of ancient (1.8×10^9 yr old) natural nuclear reactors in the uranium deposits of Oklo (Gabon, West Africa)^{9,10} have proved to be important in this respect.

The isotopic composition of Sm and Eu has been measured¹¹ for samples in the reactor core, irradiated by an independently determined¹² integrated flux of thermal neutrons $\Phi t \approx 10^{21}$ neutrons cm^{-2} . Given Φt and fission yields one can determine the capture cross-section values $\approx 1.8 \times 10^9$ yr ago. Three standard deviations give the possible range of the cross-section variation, which is connected with the resonance energy shift through the Breit-Wigner formula. One is thus led to the restriction $|\Delta E_{\text{exp}}| \lesssim 0.05$ eV, and to the estimates of the upper bounds on the variation of fundamental nuclear constants shown in Table 1 along with the earlier limits of Dyson and Davies.

These estimates seem to exclude all variants of nuclear constants change based on Dirac's 'Large Numbers Hypothesis'¹. It is, however, desirable to obtain as strict bounds on ΔE_{exp} as possible. Precise measurements of the isotopic shifts for all rare-earth fission products in the reactor core are desirable in this respect.

I would like to express my gratitude to V. E. Bunakov, V. N. Efimov, A. N. Erykalov, V. A. Ruban, and especially to Yu. V. Petrov for discussions and comments.

A. I. SHLYAKHTER

Theory Department,
Leningrad Nuclear Physics Institute,
Gatchina Leningrad District,
188350 USSR

Received May 19; accepted September 3, 1976.

- 1 Dirac, P. A. M., *Nature*, **139**, 323 (1937).
- 2 Gamow, G., *Phys. Rev. Lett.*, **19**, 759 (1967).
- 3 Dyson, F. J., *Phys. Rev. Lett.*, **19**, 1291 (1967).
- 4 Dyson, F. J., in *Aspects of Quantum Theory* (edit. by Salam, A., and Wigner, E.), 213 (Cambridge University Press, London, 1972).
- 5 Davies, P. C. W., *J. Phys.*, **A5**, 1296 (1972).
- 6 Brookhaven National Laboratory, *BNL-325*, 1 (1973).
- 7 Bohr, A., and Mottelson, B., *Nuclear Structure*, vol. 1 (Addison Wesley, New York, 1969).
- 8 Shapiro, I. S., *Usp. fiz. Nauk*, **95**, 647 (1968).
- 9 Naudet, R., *Bilans des études au 1-er mai 1974*, *Bull. Inf. Sci. Tech. (Paris)*, **193** (1974).
- 10 *The Oklo Phenomenon*, *Proc. Symp.*, 23-27 June, 1975 (IAEA, Vienna, 1975).
- 11 Dozol, J.-F., and Neuilly, M., *ibid.* p. 357.
- 12 Neuilly, M., and Naudet, R., *ibid.* p. 541.

Why measure astrophysical X-ray spectra?

MANY of the interesting results of X-ray astronomy such as the presence of compact sources in close binary systems, have been derived from light curve studies¹, obtained with quite simple detectors. On the other hand, a high resolution spectrometer, one of the most sophisticated pieces of instrumentation, is almost invariably included in solar X-ray satellites, and is used increasingly in cosmic studies². Here we wish to stress that even spectra of the highest resolution are of limited applicability in many important astrophysical problems and also perhaps to indicate the value of cost-effective planning of expensive instrumentation in general.

In addition to providing useful data through measurements of Doppler shifts and line profiles, a prime aim of high resolution spectrometry is the inference of source structure in terms of

temperature distribution of thermal particles³⁻⁶ or energy distribution of non-thermal particles⁶. These distributions are fundamental to our understanding of source mechanisms since they define both the total energy requirements⁷ and the role of energy transfer mechanisms such as conduction⁸. In this respect the X-ray spectrum problem is equivalent to optical/ultra-violet spectral studies of the solar (and stellar) chromosphere-corona structure. The thermal problem is that of the emission from an optically thin non-isothermal atmosphere and may be expressed as^{3,5,7,8}

$$I(\epsilon) = \int_T \xi(T) F(\epsilon, T) dT \quad (1)$$

where $I(\epsilon)$ is the X-ray flux at energy ϵ , $F(\epsilon, T)$ is the emission function at a temperature T per unit emission measure of a plasma of cosmic abundances⁴, and $\xi(T)$ is the emission measure of the source, differential in temperature and normalised for the source distance.

We emphasise at the outset that this equation, and its limitations discussed below, is common to spectrometry of all plasma—solar, cosmic and laboratory.

Mathematically we have to find $\xi(T)$ from $I(\epsilon)$ by inversion of equation (1) which is a Fredholm integral equation of the first kind. Such problems are mathematically ill posed¹⁰. This is distressing since it means that very small changes in $I(\epsilon)$ produce very large changes in ξ . Alternatively, a vast range of ξ 's is capable of reproducing the spectrum within arbitrarily small observational errors.

In practice of course, $I(\epsilon)$ can only be specified at a finite number of points, say n lines or continuum intervals. Given the data vector $Y = \{I(\epsilon_i)\}$ the simplest approach then seeks to find a discrete approximation X to $\xi(T)$ which satisfies

$$AX = Y \quad (2)$$

with $X = \{\xi(T_j)\Delta T_j\}$ where the T_j ($j = 1-n$) are a suitably chosen temperature grid and $A = \{F(\epsilon_i, T_j)\}$ is the matrix representation of the emission function. The required solution can be found by direct inversion

$$X = A^{-1}Y \quad (3)$$

To illustrate the practicalities of this apparently simple situation we now consider the analysis of solar active regions (AR). Attempts to use equation (3) directly for AR analysis using five or more X-ray resonance lines have run several authors^{11,13} into difficulties with nonsensical solutions, that is negative emission measures. Disregarding this salutary warning sign, 'analyses' have been tried to elude the inversion problem by fitting a trial function for X to the data. In all published AR (and solar flare) models^{11,13-17} the fitting function has been chosen purely for expediency, without any physical foundation. In addition, the results are invariably published without proper error analysis and without due consideration of the uniqueness of the model. We now examine these points in detail for a typical active region computation.

Figure 1(a) shows an AR model of the best-fit type¹¹. (This might equally be a cosmic X-ray source model (ref. 12). Using equation (2) we evaluate the fluxes, Y , to be expected in the 5 lines on which this model was based. To test the accuracy (stability) with which X can be derived from these fluxes we perturb the elements of Y with up to 5% random errors so generating a new set $\hat{Y}$ of fluxes, and reinvert by equation (3) to obtain the corresponding new derived structure $\hat{X} = A^{-1}\hat{Y}$. Figure 1b-d show three random realisations of this procedure, and it can be seen that none of these remotely resembles the true structure Fig. 1a. The essential fact is, however, that any of the structures in Fig. 1 (and innumerable others) would reproduce the observed X-ray spectrum to better

than 5%—well within the data accuracy. Had models^{11,13} been published with such an error analysis it would have been immediately clear that their 'best fit' structures were so non-unique as to be physically useless. It follows likewise that models based on minimum χ^2 or other fitting procedures with totally empirical fitting functions containing several adjustable parameters¹⁴⁻¹⁷, are physically meaningless, since an extremely wide range of functions would yield an excellent fit to the data.

Physically this analysis means that the emitted spectrum has a very low sensitivity to any features of source temperature structure other than its crudest nature—for example mean

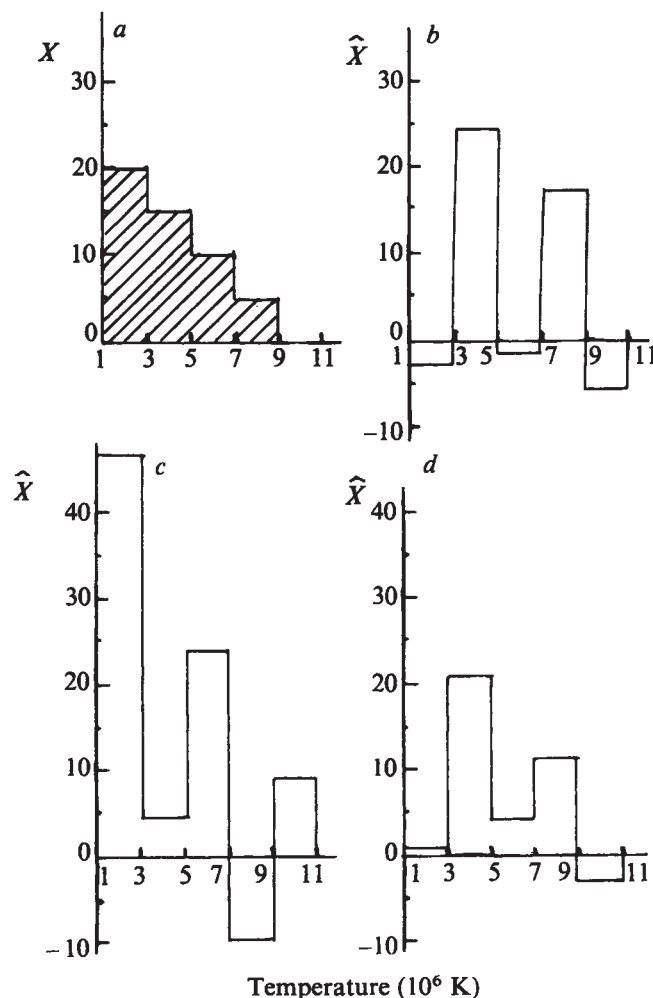


Fig. 1 a, A model (discretised) source temperature structure $X(T)$ which has a spectrum comparable to that observed in ref. 11 and satisfying the constraints they put on $X(T)$. b, c, d, Random realisations of the solutions X obtained on randomly-perturbing the spectral data by up to 5%.

temperature and variance. Theoretically, this result could have been anticipated long before the launch of any instruments. The essence of the problem is that the kernel function F in equation (1) is very flat in temperature (for both lines and continuum) which always leads to severe instability in equations of this type. Alternatively, in matrix form, equation (2), it is readily shown⁸ that errors in $\hat{Y}$ can be magnified in the solution $\hat{X}$ by factors up to η , the condition number of A . (For infinite temperature sensitivity in F , A is diagonal and $\eta = 1$, while for zero sensitivity η is infinite.) For the 5 line analysis quoted above we found $\eta \approx 30$ which explains the nonsensical solutions in Fig. 1 since the errors in $\hat{X}$ can be as large as $30 \times 5\% \approx 150\%$. It should be emphasised that this difficulty cannot be overcome merely by choice of different spectrum lines since all line responses, being governed by excitation from the

Maxwellian electron distribution, contain the factor $\exp(-\varepsilon/kT)$ which is the root of the ill-conditioning of equation (1).

Although the stability problem discussed above is well recognised in other branches of physical science^{18,19} it has so far been neglected in the field of X-ray astronomy. Here, however, the inversion problem is particularly important since knowledge of plasma temperature structure has a direct bearing on other applications of X-ray spectra. For example, computations of element abundance (and also oscillator strengths) are always based on some assumption relating the temperature structure of the source to the temperature sensitivity of lines¹⁷. Thus such methods cannot be decoupled from the difficulties presented here.

In conclusion we emphasise that these snags are not restricted to the inference of thermal structure for solar and cosmic plasmas; they also arise in non-thermal source analysis²⁰. While we agree that limited information on source structures can be obtained from spectral observations, our analysis indicates that such rudimentary information may well be obtained with instrumentation of much lower resolution than is currently available¹⁸. It is vital therefore, in assessing the cost-effectiveness of further X-ray instrumentation, to base future experiments on sound analysis of what is scientifically possible rather than on what is technologically feasible. We therefore recommend a much greater degree of theoretical sophistication in planning satellite experiments since the problem of ill-conditioning, and hence the potential ill-conception of experiments, is indeed universal in physical modelling from data.

We would like to acknowledge the support of a research grant from the SRC.

I. J. D. CRAIG
J. C. BROWN

Department of Astronomy,
University of Glasgow,
Glasgow, G12 8QQ, UK

Received December 17, 1975; accepted October 7, 1976.

- 1 Blumenthal, G. R., and Tucker, W. H., *A. Rev. Astr. Astrophys.*, **12**, 23-46 (1974).
- 2 Novick, R., *IAU Symp.*, **55**, 105-118 (1973).
- 3 Jeffries, J. T., Orrall, F. Q., and Zirker, J. B., *Solar Phys.*, **22**, 307-343 (1972).
- 4 Illarionov, A. F., and Sunyaev, R. A., *Astrophys. Space Sci.*, **19**, 61-74 (1973).
- 5 Brown, J. C., *IAU Symp.*, **57**, 395-412 (1974).
- 6 Brown, J. C., *Solar Phys.*, **18**, 489-502 (1971).
- 7 Craig, I. J. D., thesis, Univ. London (1975).
- 8 Craig, I. J. D., and Brown, J. C., *Astr. Astrophys.*, **49**, 239-250 (1976).
- 9 Tucker, W. H., and Koren, M., *Astrophys. J.*, **168**, 283-311 (1971).
- 10 Franklin, J. N., *J. math. Anal.*, **31**, 682-716 (1970).
- 11 Batstone, R. M., Evans, K., Parkinson, J. H., and Pounds, K. A., *Solar Phys.*, **13**, 389-400 (1970).
- 12 Fabian, A., and Pringle, J., *Nature*, **260**, 14 (1976).
- 13 Parkinson, J. H., *IAU Symp.*, **68**, 45-64 (1975).
- 14 Chambe, G., *Astr. Astrophys.*, **12**, 210-218 (1971).
- 15 Dere, K. P., Horan, D. H., and Kreplin, R. W., *Solar Phys.*, **36**, 459-472 (1974).
- 16 Acton, L. W., Catura, R. C., Mayeratt, A. J., Wolfson, C. J., and Culhane, J. L., *Solar Phys.*, **26**, 183-202 (1972).
- 17 Jordan, C., *IAU Symp.*, **68**, 109-134 (1975).
- 18 Kunasz, C. V., Jeffries, J. T., and White, O. R., *Astr. Astrophys.*, **28**, 15-26 (1973).
- 19 Phillips, D. L., *J.A.C.M.*, **9**, 84-97 (1962).
- 20 Brown, J. C., *IAU Symp.*, **68**, 245-282 (1975).

ground regions. It is clear from Fig. 1 that the flux is modulated throughout with a period of ~ 8 min. Using the power spectrum analysis technique described in ref. 6, we find a period of 7.90 ± 0.03 min, with a probability that this could arise from random noise of 10^{-17} and an average modulation amplitude of $23 \pm 5\%$ of the mean flux. Because of the finite sampling time (86.509 s) we cannot exclude the possibility that we are observing an alias of a modulation shorter than the Nyquist frequency of the data. Table 1 gives the first four periods which alias to 7.9 min,

Table 1 Possible aliases to 7.90 min

min	s	Attenuation
7.90 ± 0.05	474.0 ± 3.0	0.97
1.764 ± 0.003	105.84 ± 0.18	0.52
1.219 ± 0.002	73.14 ± 0.12	0.17
0.7933 ± 0.0005	47.598 ± 0.030	0.20
0.6606 ± 0.0004	39.637 ± 0.021	0.20

along with the expected attenuation for a sinusoidal modulation. The mean flux was $\sim 2.5 \times 10^{-9}$ erg cm⁻² s⁻¹, at least two to three times greater than that seen from 3U1727-33 by Uhuru⁵.

The Ariel V experiment C detector⁷ observed the same region of sky for 3 d beginning on March 8, 1976, and again for one day on March 24, 1976 (ref. 8). The experiment C field of view is circular with a f.w.h.m. of 5° , and contained within it were all the sources observed by Copernicus, plus MX1716-30 (ref. 9).

Data with 0.5, 20 and 32 s resolution were obtained in the 2-7 keV range⁸. Both observations have been subjected to a power spectrum analysis and no significant periodicities were detected in the range 45 min to 40 s, with an upper limit of 4% of the mean flux. We then repeated the analysis for each individual orbit of data. In only one orbit (7784, with a time resolution of 20 s) was a significant periodicity detected; this period was 73.1 ± 1.4 s with a probability of random occurrence of 10^{-13} . We note that this period is close to one of the possible aliased periods in the 1973 Copernicus data (Table 1). The mean amplitude of the modulation during this orbit was $14 \pm 3\%$ of the mean flux. Figure 2 illustrates the data of both this orbit and one

Table 2 Power law fit: $I(E) = CE^{-n} \exp(-\sigma N_H)$

Ariel V experiment C (2.1-20.3 keV)			
Start time day, March 1976	C	n	$N_H \times 10^{22}$
8.570	2.35 ± 0.20	2.05 ± 0.07	5.34 ± 0.02
8.642	2.29 ± 0.18	2.01 ± 0.07	5.33 ± 0.02
9.558	2.38 ± 0.17	2.05 ± 0.06	5.32 ± 0.02
9.629	2.50 ± 0.22	2.07 ± 0.07	5.33 ± 0.02
11.195	2.62 ± 0.22	2.10 ± 0.07	5.33 ± 0.02
11.252	2.52 ± 0.22	2.06 ± 0.08	5.33 ± 0.02
11.335	2.72 ± 0.24	2.11 ± 0.08	5.36 ± 0.02
Copernicus (2.8-8.7 keV)			
May 1976			
16.3	3.5 ± 1.8	2.7 ± 0.5	12 ± 6

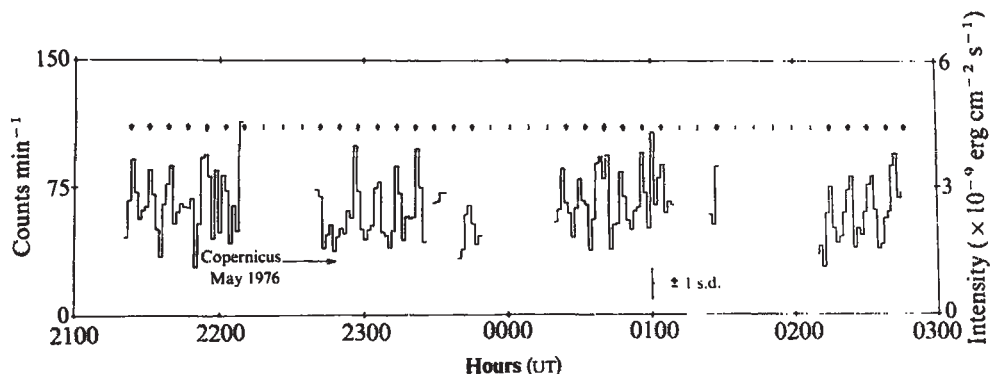
preceding it. For the orbit where a modulation was detected the predicted times of maximum flux are indicated. Upper limits on a 73-s modulation obtained for the other 20 s and also the 0.5-s orbits were typically a factor of 3 less than the amplitude of the modulation detected. The mean count rate of each orbit (a flux of $\sim 2.2 \times 10^{-9}$ erg cm⁻² s⁻¹) did not vary by more than $\sim 8\%$, in particular there was no significant difference between the mean count of the orbit 7784 and those close to it.

Spectral data were obtained on several occasions throughout the first Ariel observation with a time resolution of ~ 40 min and the results are presented in Table 2. These represent the background subtracted average of all the

Periodic behaviour of the X-ray flux from the region near 3U1727-33

THE MSSL 2.5-7.5 keV X-ray detector onboard Copernicus⁴ was pointed in the vicinity of 3U1727-33 for a total of ~ 6 h on April 14, 1973 (ref. 2). The $2.5^\circ \times 3.5^\circ$ f.w.h.m. field of view of the detector included the positions of MXB1730-335 (ref. 3), MXB1728-34 (ref. 4) and 3U1727-33 (ref. 5). A new position for 3U1727-33 has been communicated by Saulson and Forman (unpublished). The revised error box is smaller than the one reported in ref. 5 and includes the position of MXB1728-34. The background subtracted count rate observed by Copernicus has been plotted in Fig. 1. Each bin represents one 63-s accumulation followed by 24 s of dead time. The gaps in the data are caused by the occultation by the Earth of the source and passage of the satellite through high back-

Fig. 1 The count rate and flux seen in the April 14–15, 1973 Copernicus observation. The mean intensity recorded by Copernicus in May 1976 is indicated.



sources in the field of view. There is no evidence for any significant spectral variability.

A further half day observation of this region was made by Copernicus on May 16, 1976. The mean count rate observed was ~ 30 counts min^{-1} ($\sim 1.3 \times 10^{-9}$ erg cm^{-2} s^{-1}). The data as a whole show no significant modulation in the

least 50 pulses. As for the first case, there is the possibility that a repetition of the 73-s period has occurred after 3 yr. If this is so, then the modes of oscillation must be limited in number, otherwise the probability of the reappearance of a periodicity would be negligible.

We thank C. Chevalier and S. A. Ilovaisky for discussions,

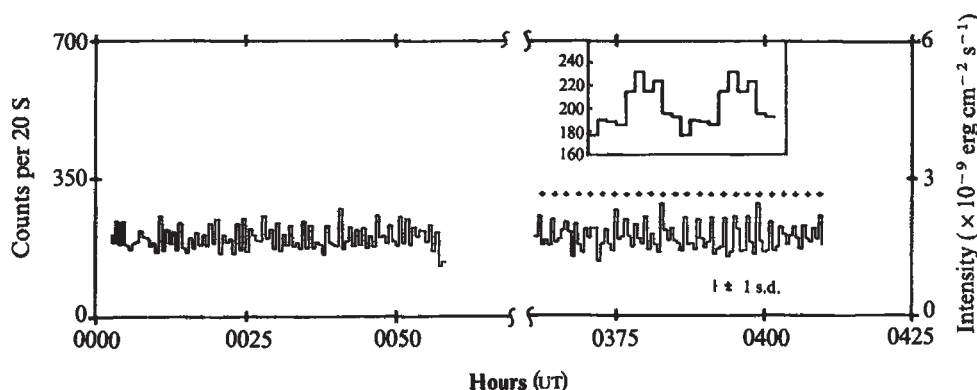


Fig. 2 The background subtracted count rate and flux observed by Ariel V experiment C during two successive orbits 7782 and 7784, in March 1976. For ease of comparison the flux axes have been set to the same as Fig. 1. Also shown are all the data from orbit 7784 folded about a period of 73.14 s and repeated once for clarity.

period 40 min–40 s with an upper limit of 12% of the mean flux. If, however, a transient periodicity such as that seen by Ariel V had been present it would not have been detected because of the poorer statistical significance of the data. Spectral information was obtained with a 6-channel PHA and the spectral parameters obtained are shown in Table 2.

There are two possible interpretations of the data:

(1) The 1973 Copernicus periodicity originated from a regularly pulsating source similar to those reported in ref. 6. Since a similar modulation was not re-observed in 1976 this suggests that either the periodicity originated from a transient source similar to that seen by Ives *et al.*¹⁰, or else the amplitude of the modulation is variable. In the latter case the 73-s periodicity seen in March 1976 could be the fundamental period of the pulsator, or its apparent relationship to the 7.9 min period is coincidental.

(2) The modulation seen in 1973 resulted from a periodic sequence of bursts originating from a burst source in the field of view. The period between bursts would be either 7.9 min or any of its aliases (Table 1). We note that such sequences of bursts have been seen from several burst sources on various time scales (refs 4, 11). In particular MXB1730–335 shows quasi-periodic bursts trains with characteristic time scales between 15 and 35 s (refs 5, 8). To explain the 1973 data we require that bursts trains can be stable for at

Mrs P. Ock for assistance with computing and Professor R. L. F. Boyd, for encouragement. The financial support of both the SRC (N.E.W. and K.O.M.) and ESA (G.B.) is acknowledged.

N. E. WHITE
K. O. MASON
G. BRANDUARDI
P. W. SANFORD

Mullard Space Science Laboratory,
University College London,
Holmbury St Mary,
Dorking, Surrey, UK

Laboratorio Fisica Cosmica,
CNR Milano, and
Istituto di Fisica dell'Universita,
Milano, Italy

L. MARASCHI
A. TREVES

Received August 16; accepted October 6, 1976.

- ¹ Sanford, P. W., *Proc. R. Soc.*, **A390**, 411 (1974).
- ² Willmore, A. P., *et al.*, *Mon. Not. R. astr. Soc.*, **169**, 7 (1974).
- ³ Lewin, W. H. G., *et al.*, *Astrophys. J. Lett.*, **207**, L95 (1976).
- ⁴ Hoffman, J. A., *et al.*, *Astrophys. J.* (in the press).
- ⁵ Giacconi, R., *et al.*, *Astrophys. J. Suppl. Ser.*, **27**, 37–64 (1974).
- ⁶ White, N. E., *et al.*, *Astrophys. J. Lett.*, **209**, L119 (1976).
- ⁷ Sanford, P. W., and Ives, J. C., *Proc. R. Soc.* (in the press).
- ⁸ Mason, K. O., Bell-Burnell, S. J., and White, N. E., *Nature*, **262**, 474 (1976).
- ⁹ Markert, T. H., *et al.*, *IAU Circ.*, No. 2765 (1975).
- ¹⁰ Ives, J. C., Sanford, P. W., and Bell-Burnell, S. J., *Nature*, **254**, 578 (1975).
- ¹¹ Clark, G. W., *et al.*, *Astrophys. J. Lett.*, **207**, L105 (1976).

Spiral structure as an explanation for the asymmetric brightness of Saturn's A ring

REITSEMA *et al.*¹ and Lumme and Irvine² have recently confirmed earlier observations which suggested that the A ring of Saturn is fainter in the quadrants following conjunctions of the particles with the Earth-Saturn line and brighter in the quadrants preceding conjunctions. Surprisingly, no intrinsic azimuthal brightness variation is found in the B ring. The brightness variation has been linked to the presence of synchronously rotating particles in the A ring, the effect caused by either a systematic variation in albedo over their surfaces or variations in their geometric projections. Of these two possibilities, the latter seems more promising since the asymmetry of the brightness variation with respect to the Earth-Saturn line is difficult to account for by means of an albedo variation. Furthermore, an albedo variation would yield the same brightness pattern for the B ring if it contained synchronously rotating particles. In this paper, however, we propose yet another mechanism to explain the phenomenon—spiral wakes in the A ring.

The asymmetry would arise in a ring of synchronously rotating, irregularly shaped, particles if there was a statistical tendency for the long axes of the particles to be inclined in the trailing sense with respect to the differential rotation. Such an orientation of the axes of the particles could arise naturally as the result of interparticle collisions and the shear from differential rotation. What is required is that the collision frequency be comparable to the libration frequency of the long axes of the particles (which is $\sim \Omega$, the orbital frequency, for irregularly shaped particles), and that the orbits of the particles deviate from circles by no more than a few times the particle radii a . The first requirement is probably met in Saturn's A and B rings, since their normal optical depths τ_1 are ~ 1 . It is the second requirement that seems unlikely to be satisfied and to which we now turn our attention.

The brightness surge of Saturn's rings as opposition is approached is most plausibly interpreted as the consequence of shadowing in rings that are many-particle radii thick. Best current estimates for the volume density give $D \sim 10^{-2}$ (ref. 3). D is simply related to the vertical thickness of the rings by $D \sim a\tau_1/h$. The random velocity of the particles $v \sim \Omega h$. These relationships imply $v \sim a\Omega\tau_1/D \sim 10^2 a\Omega$ for $D \sim 10^{-2}$ and $\tau_1 \sim 1$. Collisions will produce approximate equipartition between random and rotational kinetic energies, and thus we conclude that the particles in Saturn's rings should have spin periods that are $\sim 10^{-2}$ times their orbital periods. Of course, this conclusion applies in an average sense to the particles that scatter the sunlight. The angular spin velocities ω of individual particles should vary as $(\omega - \Omega) \propto a^{-3/2}$. Thus, particles whose radii are larger than average (the average radius $\bar{a}$ clearly being determined by weighting with respect to surface area) by a factor equal to or greater than

$$\frac{a_1}{\bar{a}} \sim 2^{1/5} \left[\frac{\tau_1}{D} \right]^{2/5}$$

would have $|\omega - \Omega|/\Omega < 1$. A significant fraction of such large particles would be at least temporarily trapped in synchronous rotation if they were irregular in shape. The azimuthal brightness variation might be accounted for if large particles, $a > a_1$, scattered a significant fraction ($\geq 10\%$) of the sunlight incident on the A ring. There is, however, a more attractive mechanism which we shall now describe that is capable of producing the observed effect.

Large particles, whose surface escape velocities are $> v$, will force intense trailing density wakes. Such wakes have been studied previously in connection with spiral structure in galaxies⁴. The detailed mechanics of the formation of wakes in a ring composed of inelastically colliding particles is certain to

be complicated. On the basis of the study cited above, however, there is no doubt that in differentially rotating media these wakes are always oriented in the trailing sense! The dimensions of the trailing density concentrations will range up to a few times the ring thickness h . Since $h \lesssim 4$ km (ref. 3), these density wakes are much too small to be resolved from the Earth but might be resolved by the Mariner Jupiter/Saturn spacecraft due to fly by the planet in 1980. The density wakes, if optically thick, will scatter sunlight in a manner that mimics scattering by synchronously rotating, irregular particles whose long axes are inclined in a trailing sense. The particles in a wake would, however, be only transient neighbours and in no way would a particle and its wake constitute a rigidly rotating entity.

Particles whose surface escape velocities are $\sim v$ have radii $a_2 \sim v/(8G\rho)^{1/2}$. Using the relationships developed previously, we find

$$\frac{a_2}{\bar{a}} \sim \frac{\Omega}{2(2G\rho)^{1/2}} \left(\frac{\tau_1}{D} \right)$$

This expression is to be compared with the expression for $a_1/\bar{a}$ derived earlier. It is easily shown that

$$\frac{a_2}{a_1} \sim 2^{-7/10} \frac{\rho_s}{\rho} \left(\frac{R_s}{r} \right)^3 \left(\frac{\tau_1}{D} \right)^{3/5}$$

where $\rho_s = 0.7$ g cm⁻³ and $R_s = 6 \times 10^9$ cm are the mean density and the mean radius of Saturn. Taking $\rho = 1$ g cm⁻³, $r = 1.3 \times 10^{10}$ cm, $\tau_1 = 0.5$ and $D = 10^{-2}$, which are appropriate parameters for ice particles near the centre of the A ring, we obtain $a_2/a_1 \sim 0.4$. Thus, particles that are large enough to rotate synchronously will also force dense trailing wakes.

To this point, our discussion has been based on the assumption that the opposition effect implies $D \sim 10^{-2}$. We stress that if this assumption were incorrect and $D \gg 10^{-2}$, our conclusion, that density fluctuations rather than synchronously rotating particles are responsible for the azimuthal brightness variation, would be strengthened (note $a_2/a_1 \propto D^{-3/5}$). Furthermore, if the mean mass density in the rings were as high as 0.1 g cm⁻³, the rings would be on the brink of gravitational instability^{5,6}. In this case, collective gravitational effects would greatly enhance the strength of trailing density perturbations⁴.

The absence of an azimuthal brightness variation in the B ring must still be explained. Two possibilities immediately come to mind. First, the higher optical depth of the B ring would make its brightness less sensitive to either the geometrical orientations or the spatial correlations of the individual particles. Second, both mechanisms for producing the azimuthal brightness variation require the rings to contain particles whose radii are considerably larger than average. Small particles are more likely to gravitationally aggregate and form large particles in the A ring than in the B ring. In this context, we note that the formal Roche limit for particles with $\rho = 1$ g cm⁻³ occurs at $r = 1.3 \times 10^{10}$ cm, near the centre of the A ring.

G. COLOMBO

Harvard Smithsonian,
Center for Astrophysics,
60 Garden Street,
Cambridge, Massachusetts 02138

P. GOLDBREICH

California Institute of Technology,
Pasadena, California 91125

A. W. HARRIS

Jet Propulsion Laboratory,
Pasadena, California 91103

Received August 25; accepted October 7, 1976.

¹ Reitsema, H. J., Beebe, R. F., and Smith, B. A., *Astr. J.*, **81**, 209–215 (1976).

- ² Lumme, K., and Irvine, W. M., *Astrophys. J. Lett.*, **204**, L55-L57 (1976).
³ Pollack, J. B., *Space Sci. Rev.*, **18**, 3-93 (1975).
⁴ Julian, W. H., and Toomre, A., *Astrophys. J.*, **146**, 810-830 (1966).
⁵ Cook, A. F., and Franklin, F. A., *Astr. J.*, **69**, 173-200 (1964).
⁶ Safronov, V. S., *Evolution of the Protoplanetary Cloud and Formation of the Earth and the Planets* (Israel Program for Scientific Translations, Jerusalem 1972).

Circumstellar acetylene in the infrared spectrum of IRC +10° 216

A THICK, expanding envelope of dust and gas surrounds the unusual object IRC +10° 216. Microwave studies have identified numerous molecular constituents of the envelope¹, and the thermal infrared spectrum indicates the probable presence of both silicon carbide and graphite in the form of dust grains (ref. 2, and M. P. Campbell, *et al.*, unpublished). We report here the detection of acetylene, C_2H_2 , in the infrared spectrum. Although acetylene is expected to be an important constituent of some stellar atmospheres, and possibly of interstellar molecular clouds, this appears to represent the first positive detection of C_2H_2 outside the solar system.

A Fourier transform spectrometer at the coude focus of

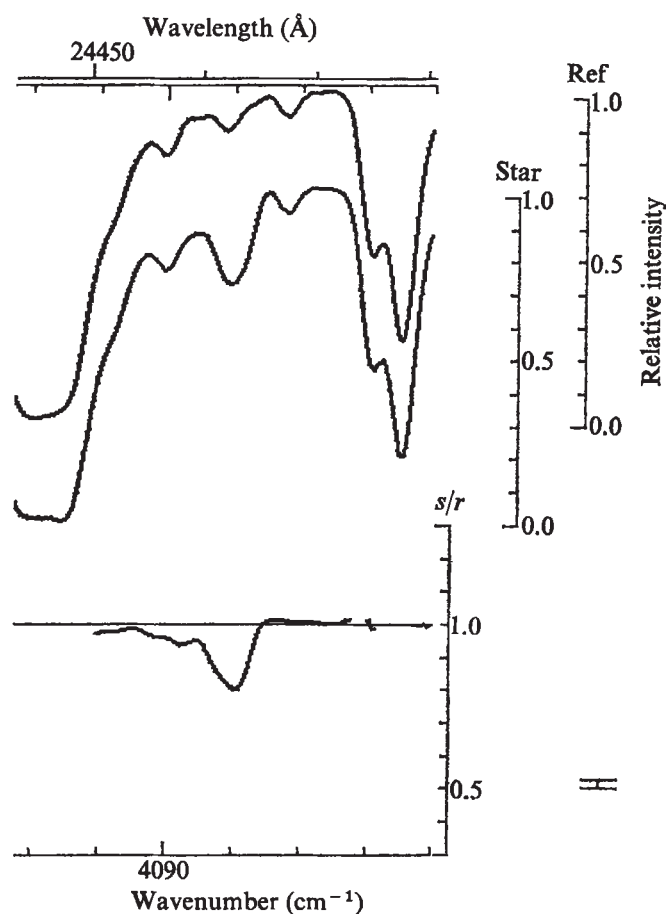


Fig. 1 The spectrum of IRC +10° 216. Above, a comparison solar spectrum (dominated by terrestrial absorption); middle, the sum of two +10° 216 spectra; below, the ratio +10° 216/sun. A 2σ error bar, computed from the difference of the two +10° 216 spectra, is shown at the lower right. The resolution is 0.25 cm^{-1} , and the FWHM of the instrumental line profile is 0.30 cm^{-1} . The abscissa scale at the bottom is in units of $1\text{ cm}^{-1}/\text{tic}$ (vacuum) increasing to the right; at the top, units are $10\text{ Å}/\text{tic}$ (STP) increasing to the left. The Doppler shift of +10° 216 at the time of observation was negligible for purposes of this figure, but has been accounted for in the analysis. In the ratio +10° 216/sun the small peak at 4090.52 cm^{-1} is an artefact arising from a weak absorption line in the solar spectrum.

the Mayall 4-m telescope at Kitt Peak was used to obtain a spectrum of the region $4,000\text{--}5,000\text{ cm}^{-1}$ ($2.0\text{--}2.5\text{ }\mu\text{m}$). This was a daytime observation, and a solar reference spectrum was recorded with a small heliostat shortly before and at similar zenith distance. A very small portion of these spectra are shown in Fig. 1. The +10° 216 spectrum contains a resolved, red degraded feature at $4,091\text{ cm}^{-1}$ which we attribute to the $\nu_1 + \nu_5$ band of acetylene³. Only the Q branch is detectable. Lines Q(1) to Q(7) fall in an interval of only 0.12 cm^{-1} . In the P and R branches, the lines are too widely separated to yield distinctive features.

The strength of the feature drops sharply near $J=16$, suggesting a temperature of $200\text{--}300\text{ K}$, so the absorption is most probably associated with the $2''$ diameter circumstellar shell^{4,5}. Adopting a temperature of 250 K , the strongest acetylene absorption would be expected to occur at $4,090.98\text{ cm}^{-1}$. The observed maximum absorption, corrected for the Doppler shift appropriate for the $2''$ shell (ref. 5 and D. N. B. Hall, *et al.*, unpublished), falls at $4,090.97 \pm 0.03\text{ cm}^{-1}$, indicating the internal consistency of our interpretation. From published spectra^{6,7} and intensity studies^{8,9} we estimate a $\nu_1 + \nu_5$ Q branch intensity of $0.2\text{ cm}^{-2}\text{ atm}^{-1}$. The observed equivalent width of 0.2 cm^{-1} then requires a column abundance $\sim 1\text{ cm atm}$, or $3 \times 10^{19}\text{ mol cm}^{-2}$. For a $2''$ diameter shell and a distance of 200 pc (D. N. B. Hall, *et al.*, unpublished) this corresponds to a total mass of $\sim 10^{-4} M_\odot$.

Of the other important C_2H_2 bands only the $3.04\text{-}\mu\text{m}$ band⁹ should definitely be detectable in existing +10° 216 spectra. A previously unidentified band in fact occurs in many carbon stars¹⁰ at the correct position. The strength of this band in +10° 216 ($\sim 90\text{ cm}^{-1}$) could reasonably arise from 1 cm atm of acetylene, and we suggest that acetylene does in fact contribute part or all of this feature. Absorption in the $3\text{ }\mu\text{m}$ region from other hydrocarbons is, however, likely in carbon-star photospheres, and cannot be ruled out in the +10° 216 shell.

Except for the $2\text{--}0$ band of CO (D. N. B. Hall, *et al.*, unpublished) no other lines were found in the $4,000\text{--}5,000\text{ cm}^{-1}$ region to a limiting equivalent width of $\sim 0.05\text{ cm}^{-1}$. From laboratory spectra¹¹ of the 2.32 and $2.37\text{ }\mu\text{m}$ bands of methane, we estimate a CH_4 abundance $< 0.03\text{ cm atm}$, or $7 \times 10^{17}\text{ mol cm}^{-2}$.

IRC +10° 216 is apparently a site of continuing production and ejection of grains and molecules. In view of the serious difficulties encountered in developing a satisfactory understanding of these processes, +10° 216 deserves detailed and continuing study. In particular, two important questions may be resolved by continued investigation: first, is grain formation a cause or an effect of mass ejection?; second, do molecular abundances 'freeze out' on ejection? Our results for C_2H_2 and CH_4 (which are not detectable in the microwave region because of lack of pure rotational dipole transitions) are particularly relevant to the second question. The large C_2H_2 abundance, comparable to the CO abundance (ref. 5 and D. N. B. Hall, *et al.*, unpublished), is consistent with computations for cool, carbon rich stellar photospheres¹². The predominance of C_2H_2 over CH_4 is also expected. If the photosphere were ejected in such conditions that it cooled while maintaining molecular equilibrium, however, C_2H_2 should be strongly depleted¹³. Morris¹ has compared relative abundances of SiS, SiO, CS and HC_3N with recent molecular equilibrium calculations¹⁴. While inconclusive, his results did not rule out 'freezing' of photospheric abundances. Additional spectroscopic study should provide crucial relative abundances (or upper limits) for C_2 , CN, HCN, and possibly others, permitting a more conclusive analysis of the evolution of the molecular abundances during ejection.

A number of infrared objects are known which appear to be similar to +10° 216 (ref. 10). In addition, certain bright carbon stars show a $10\text{-}\mu\text{m}$ flux excess, attributed to circum-

stellar dust¹⁵. We have obtained a spectrum of one of the latter, V Cygni. In this star the 4,091 cm⁻¹ band appears prominently, stronger than in +10° 216, and apparently originating in a circumstellar region of somewhat higher temperature. It is possible that the circumstellar acetylene and dust grains are mixed at a common kinetic temperature. If so, progress may be possible in the difficult problems of determining grain emissivity and particle size distribution.

STEPHEN T. RIDGWAY
DONALD N. B. HALL

Kitt Peak National Observatory,
Tucson, Arizona

SUSAN G. KLEINMANN
DOREEN A. WEINBERGER

Massachusetts Institute of Technology,
Cambridge, Massachusetts

ROBERT S. WOJSLAW

Kitt Peak National Observatory,
Tucson, Arizona

Received August 24; accepted October 15, 1976.

- ¹ Morris, M., *Astrophys. J.*, **197**, 603–610 (1975).
- ² Treffers, R., and Cohen, M., *Astrophys. J.*, **188**, 545–552 (1974).
- ³ Baldacci, A., Gherse, S., and Rao, K. N., *J. molec. Spectrosc.*, **48**, 600–603 (1973).
- ⁴ Toombs, R. I., et al., *Astrophys. J. Lett.*, **173**, L71–L74 (1972).
- ⁵ Geballe, T. R., Wollman, E. R., and Rank, D. M., *Astrophys. J.*, **183**, 499–504 (1973).
- ⁶ Cruikshank, D. P., *Comm. lunar planet. Lab.*, **79**, 1–30 (1967).
- ⁷ Pierson, R. H., Fletcher, A. N., and Gantz, E. S., *Nat. Bur. Stand. J. of Res.*, **28**, 1218–1239 (1956).
- ⁸ Varanasi, P., and Bangaru, B. R. P., *J. Q.S.R.T.*, **14**, 839–844 (1974).
- ⁹ Varanasi, P., and Bangaru, B. R. P., *J. Q.S.R.T.*, **15**, 267–273 (1975).
- ¹⁰ Merrill, K. M., and Stein, W. A., *Pub. astr. Soc. Pacif.*, **88**, 294–307 (1976).
- ¹¹ Kuiper, G. P., and Cruikshank, D. P., *Commun. lunar planet. Lab.*, **34**, 141–165 (1964).
- ¹² Tsuji, T., *Ann. Tokyo astr. Obs.*, 2nd Ser., **9**, 1–110 (1964).
- ¹³ Salpeter, E. E., *Astrophys. J.*, **193**, 579–584 (1974).
- ¹⁴ Tsuji, T., *Astr. Astrophys.*, **23**, 411–431 (1973).
- ¹⁵ Gillett, F. C., Merrill, K. M., and Stein, W. A., *Astrophys. J.*, **164**, 83–90 (1971).

Time-varying Newtonian gravity and universal motion

A time-dependent gravitational constant $G(t)$ was proposed by Dirac¹ in 1937, and there is some limited evidence for it^{2–4}. It can be incorporated in relativistic theories of gravitation (see refs 5–7), but the resulting equations are liable to be cumbersome, and it is uncertain which theory should be used. Attempts to consider this problem in a Newtonian framework^{3,4,8–10} have been based on the amended Newtonian force law

$$\mathbf{F} = -G(t) \frac{m_1 m_2 \mathbf{r}}{r^3} \quad (1)$$

We show here that the assumption of energy conservation (as discussed below) leads to a more substantial change in the force law.

That equation (1) implies violation of energy conservation for particle interactions can be seen as follows. One particle, shaped like a doughnut, moves from infinity directly towards a second particle; the particles move through each other to infinity. If $G(t)$ is decreasing, then, for the same distance of separation, the force of attraction whilst the particles are approaching is greater than the force whilst the particles are moving apart. The final relative velocity for large separation is greater than the initial relative velocity. Thus the final kinetic energy is greater than the initial kinetic energy. Since the potential energy is zero for large separation, energy conservation for interacting particles has been violated; the energy of the gravitational field is

not involved. Energy conservation for interacting static bodies would of course also be violated.

We see that energy conservation and equation (1) cannot both hold in the generalised Newtonian theory. In the past the failure of the energy law appears to have been ignored, and equation (1) has been maintained, presumably as the more physically important of the two laws. In this note an alternative path is explored: a new force law will be obtained from energy conservation, which, together with expressions for kinetic and potential energy, will be assumed.

We consider a finite set of discrete particles with masses m_α ($\alpha = 1, \dots, N$). We shall take the potential energy, V , of the system in the general form

$$V = \frac{1}{2} \sum_{\substack{\alpha, \beta=1 \\ \alpha \neq \beta}}^N m_\alpha m_\beta v(r_{\alpha\beta}, t) \quad (2)$$

Referring to a chosen inertial frame, $\mathbf{x}_\alpha$ is the position vector of the particle α ; $\mathbf{r}_{\alpha\beta} = \mathbf{x}_\alpha - \mathbf{x}_\beta$; $r_{\alpha\beta} = |\mathbf{r}_{\alpha\beta}|$; and the function

$$v(r_{\alpha\beta}, t) = \frac{-G(t)}{r_{\alpha\beta}} \quad (3)$$

would be the obvious choice. The kinetic energy of the system has the usual form, and the energy conservation law is

$$\sum_{\alpha=1}^N \frac{1}{2} m_\alpha \sum_{i=1}^3 \dot{x}_{\alpha i}^2 + \frac{1}{2} \sum_{\substack{\alpha, \beta=1 \\ \alpha \neq \beta}}^N m_\alpha m_\beta v(r_{\alpha\beta}, t) = E \quad (4)$$

where E is a constant. Total differentiation with respect to time of equation (4) leads, after some algebra, to

$$0 = \sum_{\alpha=1}^N m_\alpha \sum_{i=1}^3 \dot{x}_{\alpha i} \left(\ddot{x}_{\alpha i} + \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^N m_\beta \left[\frac{\partial v}{\partial r_{\alpha\beta i}} + \frac{\dot{r}_{\alpha\beta i}}{V_{\alpha\beta}^2} \frac{\partial v}{\partial t} \right] \right) \quad (5)$$

where $V_{\alpha\beta} = |\dot{\mathbf{r}}_{\alpha\beta}|$. Equation (5) suggests, and is consistent with, the force law ($i = 1, 2, 3$)

$$\ddot{x}_{\alpha i} = - \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^N m_\beta \left[\frac{\partial v}{\partial r_{\alpha\beta i}} + \frac{\dot{r}_{\alpha\beta i}}{V_{\alpha\beta}^2} \frac{\partial v}{\partial t} \right] \quad (6)$$

A formal approach to the preceding paragraph is as follows. Newton's three laws of motion apply, and the particle interaction force is assumed to be given by equation (6). Equation (5) follows immediately, and may be integrated to show that equation (4) is a constant of the motion. Because of the form of equation (4), it is physically appealing to identify this constant as the energy.

The ratio, X , of the second term to the first term in equation (6), subject to equation (3), is estimated for a particle moving with speed V (cm s⁻¹) on the surface of the Earth (taken as a sphere of radius r)

$$|X| \simeq \left| \frac{\dot{G}}{G} \right| \frac{r}{V} < \frac{10^{-8}}{V}$$

where $|\dot{G}/G| < 10^{-10} \text{ yr}^{-1}$ has been used. For two atoms at a distance apart, a , $\sim 10^{-8} \text{ cm}$ in a solid at absolute temperature T , taking $V \simeq (kT/m)^{1/2}$, one finds similarly

$$|X| \approx \left(\frac{m}{kT} \right)^{1/2} a \left| \frac{\dot{G}}{G} \right| < \frac{10^{-29}}{T^{1/2}}$$

The correction to the second term in equation (6) is thus normally negligible.

Even if equation (3) is not used, the force law, equation (6), has the remarkable implication that if the law of gravitation, equation (2), is time-dependent, then, to avoid infinite forces, no two particles in the Universe may be at relative rest for a non-zero time interval. The subject of statics must be considered as an approximation if the above ideas are accepted, quite apart from arguments based on quantum mechanics. Our principle of universal motion is in agreement with the observed fact that physical systems on all scales are in states of relative motion: there are zero point vibrations in condensed systems; planets orbit stars; galaxies rotate; and the Universe itself is expanding.

In conclusion, we note the following points:

(1) Newtonian cosmology with time-varying gravity has been discussed previously using an energy conservation law^{11,12}. The application of the force law, equation (6), to cosmology has been considered¹³, and yields results consistent with the energy approach.

(2) Equation (6) is Machian, in that the force on a particle depends on the velocity of all other particles in the Universe.

(3) Since equation (6) leads to an interparticle force which need not act along the line joining the two particles, angular momentum is not conserved in general.

(4) It would be desirable to find a variational principle leading to equation (6).

One of us (N.T.B.) is indebted to the SRC for support.

N. T. BISHOP*
P. T. LANDSBERG

Department of Mathematics,
The University,
Southampton SO9 5NH, UK

Received August 27; accepted October 19, 1976.

* Present address: Department of Applied Mathematics, The University of the Witwatersrand, Jan Smuts Avenue, Johannesburg, South Africa.

- ¹ Dirac, P. A. M., *Nature*, **139**, 323 (1937).
- ² Wesson, P. S., *Q. J. R. astr. Soc.*, **14**, 9 (1973).
- ³ van Flandern, T. C., *Mon. Not. R. astr. Soc.*, **170**, 333 (1975).
- ⁴ Dearborn, D. S., and Schramm, D. N., *Nature*, **247**, 441 (1974).
- ⁵ Brans, C., and Dicke, R. H., *Phys. Rev.*, **124**, 925 (1961).
- ⁶ Ni, W.-T., *Astrophys. J.*, **176**, 769 (1972).
- ⁷ Hoyle, F., and Narlikar, J. V., *Action at a Distance in Physics and Cosmology* (Freeman, San Francisco, 1974).
- ⁸ Vinti, J. P., *Mon. Not. R. astr. Soc.*, **169**, 417 (1974).
- ⁹ Roxburgh, I. W., *Nature*, **261**, 301 (1976).
- ¹⁰ Lewis, B. M., *Nature*, **261**, 302 (1976).
- ¹¹ Landsberg, P. T., *Nature Phys. Sci.*, **244**, 66 (1973).
- ¹² Landsberg, P. T., and Bishop, N. T., *Mon. Not. R. astr. Soc.*, **171**, 279 (1975).
- ¹³ Bishop, N. T., thesis, Univ. Southampton (1976).

Thermal alteration of young kerogen in relation to petroleum genesis

AVAILABLE evidence indicates that petroleum hydrocarbons can be generated by alteration of kerogens through geothermal heating¹⁻³. In this study we have separated kerogen, humic acid and lipid (benzene-methanol extractable) material from a young marine sediment (Tanner Basin, offshore California) and heated it in sealed glass tubes in a nitrogen atmosphere. Each tube was heated at a single temperature between 150 and 410 °C for a specified time in the range 5-120 h. Gaseous and liquid (benzene-methanol extractable) products generated during the heating, as well as the residual organic material, were characterised by gas-liquid chromatography, elemental analysis, infrared and electron

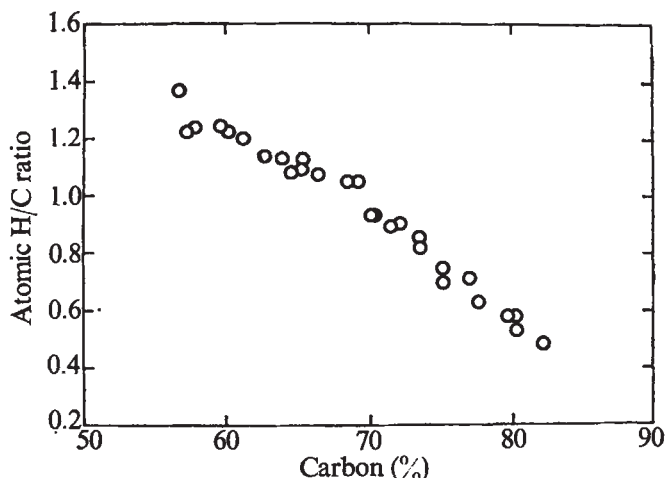


Fig. 1 Changes in H/C ratio and percentage C of kerogen when heated.

spin resonance (ESR) spectroscopies, and X-ray diffraction.

The sediment sample used has approximately 6% (by dry weight) of organic carbon. The lipid, humic acid and kerogen fractions account for 10, 7 and 72% of the total organic matter, respectively. Unheated kerogen contains 56.7% carbon, 6.5% hydrogen, 6.0% nitrogen and 30.8%

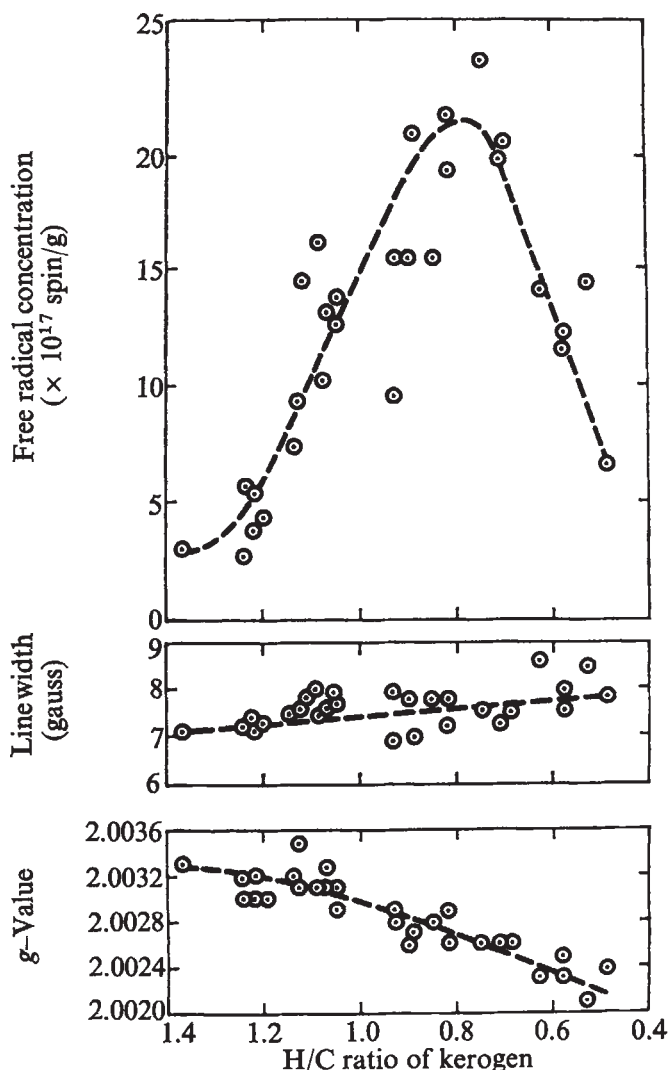


Fig. 2 Relationship between ESR properties of kerogen and its H/C ratio.

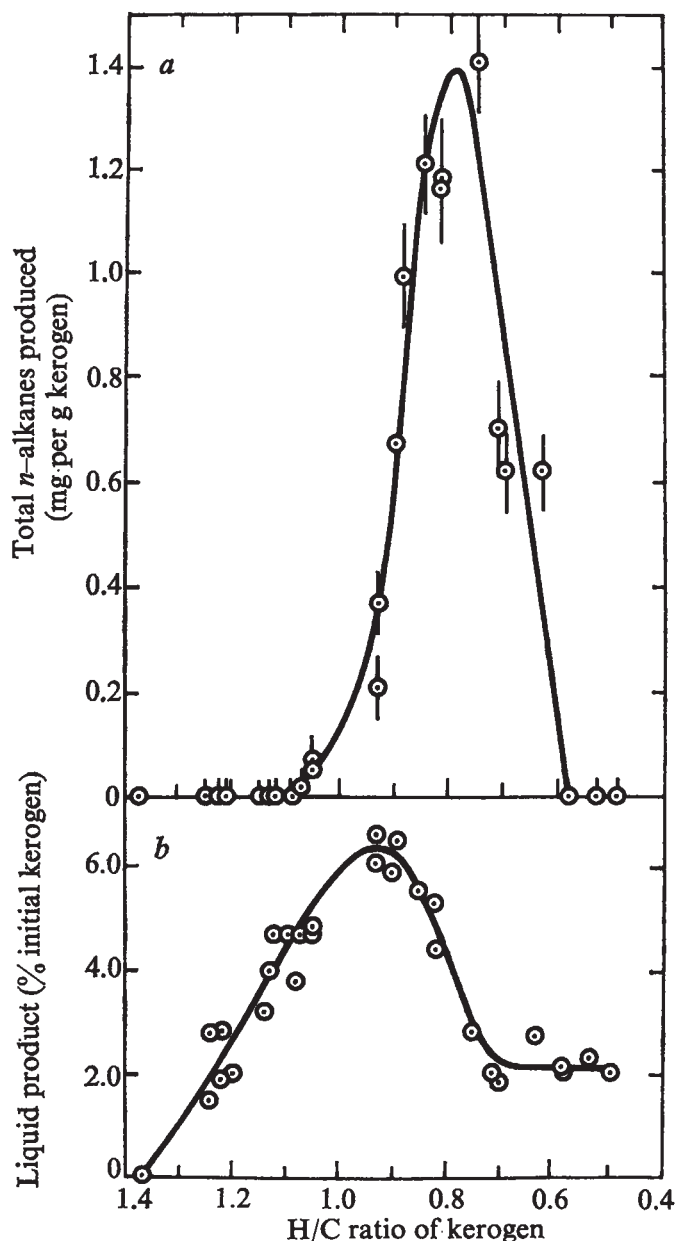


Fig. 3 Total normal alkanes and liquid product generated are plotted against the H/C ratio of kerogen.

oxygen (by difference) on an ash-free basis, and seems to fit into the type III kerogen classification of Tissot *et al.*³.

With increasing time and temperature, between 150 and 410 °C, the carbon content of kerogen increased from 56.7 to 82.1% and its atomic H/C ratio decreased from 1.37 to 0.49 (Fig. 1). Therefore, the values of either percentage C or the H/C ratio of residual kerogen may be used as an index for "level of organic maturation" (LOM) for this type of kerogen. Figure 2 shows plots of H/C ratios of residual kerogen against various parameters derived from ESR spectroscopy. With increasing degree of thermal alteration of kerogen, spin concentration increases, which may be closely associated with bond cleavage and hence, the elimination of functional groups. Below an H/C value of 0.7, the spin concentration decreases again, due to the gradual conversion of kerogen to a graphite-like structure, as verified by X-ray diffraction analysis. The kerogen *g*-value decreases with increasing degree of thermal alteration. This indicates the change of free radicals from semi-quinone type to radicals of carbon, nitrogen and sulphur due to imperfection in the kerogen structure⁴. Changes in ESR

characteristics on heating follow the same general pattern as those described by Pusey⁵ for kerogen in sedimentary rocks and Van Krevelen⁶ for coals (Fig. 2).

During heating, kerogen produced large amounts of gaseous product, amounting to about 40% of the initial weight at 410 °C. In the initial stages of thermal alteration, only H₂O and CO₂ were evolved. With increasing degree of thermal alteration, hydrogen, methane and other volatile hydrocarbons began to be generated, and their content increased gradually.

In addition, liquid products soluble in organic solvents began to be produced at a low temperature of alteration, reached a maximum at a kerogen H/C ratio of 0.90–0.95, and then decreased as thermal alteration increased. The decrease in the amount of liquid product may be explained by its further decomposition on heating. In Fig. 3a, the relationship between the total (C₁₃–C₃₁) normal alkanes or paraffins generated and H/C of kerogen is recorded. For the Tanner Basin kerogen, paraffins begin to be generated around H/C=1.1, increase sharply, reach a maximum

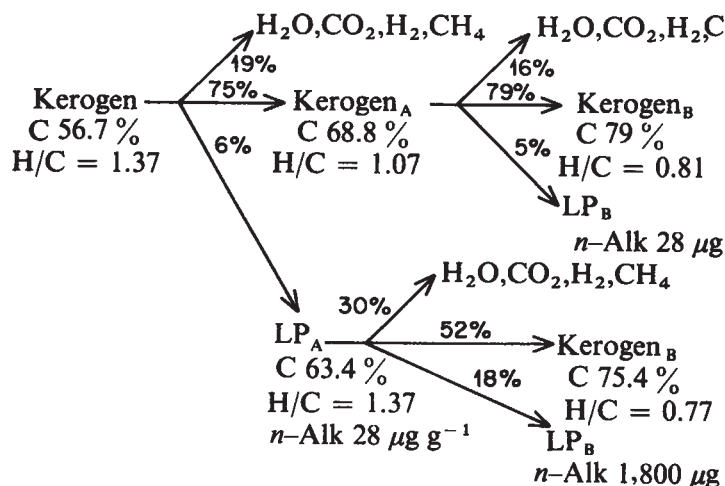


Fig. 4 Generation mechanism of *n*-alkanes from kerogen on heating.

around H/C=0.8, and then decrease rapidly. The apparent activation energy of the generation of the total *n*-alkanes was calculated to be 31 kcalorie mol⁻¹, which is lower than that of the breakage of C–C bonds in *n*-alkanes (58 kcalorie mol⁻¹). Therefore, generation of *n*-alkanes from kerogen may be due to the breakage of bonds other than C–C, for example, functional side chains on aromatic or cyclic rings, or compounds containing O, N or S, where the activation energy for bond breakage is lower than for paraffins⁸.

It is quite interesting to note that the H/C ratio of Tanner Basin kerogen at the maximum production of total liquid product (H/C=0.90–0.95; see Fig. 3b) is different from that at the maximum production of *n*-alkanes (H/C=0.8). This suggests that paraffins may be generated by the subsequent decomposition of the liquid product. A similar conclusion was reached by Tissot's group in the heating experiment of kerogen from a shale⁹. To elucidate further the mechanism of hydrocarbon generation, our kerogen was heated at 260 °C for 18.5 h, the optimum conditions for the generation of liquid product but too low to produce much *n*-alkanes. By this reaction, liquid product (LP_A) and kerogen (A) were obtained, and then those two fractions were heated again at higher temperature (330 °C, 24 h) to produce paraffins (Fig. 4). Clearly, the liquid product obtained on heating at the lower temperature is more aliphatic (H/C=1.37) than the residual kerogen (A) (H/C=1.07). By the second heating at the higher temperature, the liquid product (LP_A) produced a large amount of normal paraffins, as well as highly aromatic kerogen-like material. Heating the kerogen (A), also pro-

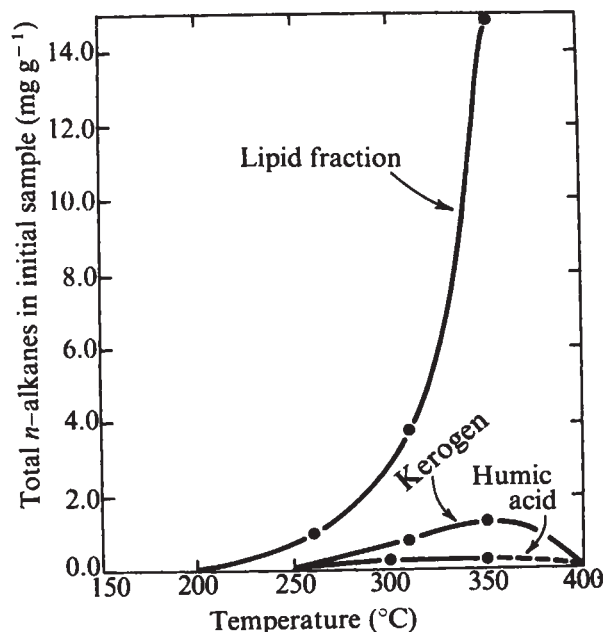


Fig. 5 Changes in the amount of total *n*-alkanes generated from various organic fractions on heating at various temperatures for 24 h.

duced alkanes and highly aromatic kerogen (B), but only about a third of the amount produced from the liquid product (LP₁). These results support the model for the production of normal hydrocarbons proposed by Tissot's group.

Interestingly, the distribution pattern of *n*-alkanes generated from kerogen during the initial stage of thermal alteration, showed an even carbon number predominance (CPI = ~0.5). With increasing degree of thermal alteration, the distribution pattern approached that of petroleum hydrocarbons (~1.0).

The results of heating other organic sediment fractions (lipid and humic acid) showed that these fractions also have an ability to produce paraffinic hydrocarbons (Fig. 5). The amounts produced are, however, very different. The contribution of each organic fraction to the total alteration was calculated for the Tanner Basin sediment sample. The lipid fraction accounts for 58%, humic acid 0.4% and kerogen 41% of the total amount of normal hydrocarbons produced.

This work was supported by a grant from NASA and by a contract from the ERDA.

R. ISHIWATARI*
M. ISHIWATARI†
I. R. KAPLAN
B. G. ROHRBACK

Institute of Geophysics and Planetary Physics,
University of California,
Los Angeles, California 90024

Received July 12; accepted September 29, 1976.

*Present address: Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Fukazawa Setagaya-ku, Tokyo, Japan.

†Present address: Engineering Research Institute, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo, Japan.

- 1 Welte, D. H., *J. Geochem. Explor.*, **1**, 117 (1972).
- 2 Tissot, B., Califet-Debyser, Y., Deroo, G., and Oudin, J. L., *Bull. Am. Ass. petrol. Geol.*, **55**, 2177 (1971).
- 3 Tissot, B., Durand, B., Espitalie J., and Combaz, A., *Bull. Am. Ass. petrol. Geol.*, **58**, 499 (1974).
- 4 Yen, T. F., Erdman, J. G., and Saraceno, A. J., *Analyt. Chem.*, **34**, 694 (1962).
- 5 Pusey, W. C., *Trans. Gulf Coast Assoc. Geol. Soc. 23rd Ann. Convention*, 195 (Houston, 1973).
- 6 Van Kervelen, D. W., *Coal—Topology Chemistry Physics Constitution* (Elsevier, New York and Amsterdam, 1961).
- 7 McNab, J. G., Smith, P. V., Jr, and Betts, R. L., *Ind. Eng. Chem.*, **44**, 2556 (1952).
- 8 Madorsky, S. J., *Thermal Degradation of Organic Polymers*, (Wiley Interscience, New York, 1964).
- 9 Louis, M. C., and Tissot, B., *Symposium 1, Proc. 7th Wild Petrol. Cong. (Mexico)*, **2**, 47 (Elsevier, Amsterdam, London, New York, 1967).

Photocell using covalently-bound dyes on semiconductor surfaces

WE have developed a new type of electrochemical photocell, using an optically transparent semiconductor electrode whose surface was modified by the chemical binding of sensitising dyes. In spite of the absence of dye in the solution, a photocurrent, as large as those previously observed at the unmodified electrodes in contact with the dye solutions was obtained.

Since the initial report by Fujishima and Honda¹ on a photocell employing a TiO₂ n-type semiconductor as the photoanode to decompose water under illumination by photons of energy greater than the band gap, there has been much research²⁻⁶ on this subject because of its important implications for solar energy conversion and hydrogen production. Photosynthetic light conversion⁷⁻¹⁰ in green plants has also been studied. From the commercial use of electrochemical photocells as solar energy conversion devices, the photoelectrodes should be stable on illumination in aqueous solutions. TiO₂ and SnO₂ are the two candidates that have been so far found to be inert¹¹⁻¹³. The band gaps of TiO₂ and SnO₂ are, however 3.0 eV (ref. 14) and 3.5 eV (ref. 15) respectively which correspond to photon wavelengths of 415 and 355 nm. Solar energy reaching the Earth's surface has a spectrum distribution in longer wavelength¹², and cannot therefore be used. Effective ways of extending the sensitivity of electrochemical photocells to longer wavelengths are therefore strongly desired. A better way to solve the problem would be spectral sensitisation, in analogy with the primary light reaction of photosynthesis, where chlorophyll molecules accept the solar energy. A noticeable anodic current can be seen in electrochemical cells containing an n-type semiconductor anode in contact with certain dye solutions when illuminated with light in the spectral region where the dye absorbs, even though of energy less than the band gap of the semiconductor. This electrochemical spectral sensitisation works for various electrode materials, including TiO₂ (ref. 16) and SnO₂ (ref. 17), and has been mainly studied to elucidate the mechanism of the spectral sensitisation, since Gerischer *et al.*¹⁸ started using ZnO single crystal.

As sensitisers, dyes belonging to the xanthene and the cyanine groups have been widely used. Gerischer and Tributsch¹⁹ observed also an increase in the sensitised photocurrent on adding of reducing agent, such as hydroquinone,

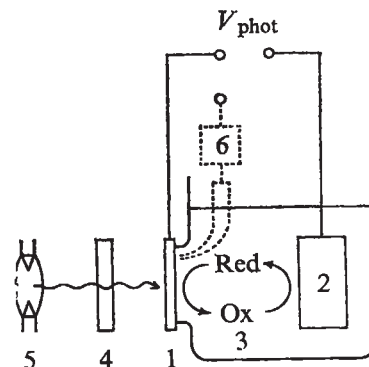
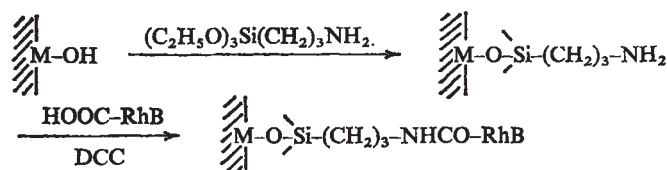


Fig. 1 An electrochemical photocell consisting of an optically transparent semiconductor anode chemically modified with rhodamine B (1) and a Pt cathode (2) in contact with an electrolyte containing a redox system (3). As a light source, a 500-W xenon lamp (5) was used in conjunction with filters and/or a monochromator (4). For measurements of spectral sensitisation (shown in Fig. 2) the electrode was polarised with a conventional potentiostat using a saturated calomel electrode (6) with a Luggin capillary.

Na_2SO_3 , or H_2O_2 , to the dye solution, although the reducing agents show no absorption in the visible region. This phenomenon is termed 'supersensitisation.'

For a real photocell, however, a new problem may arise: most of the solar energy is absorbed ineffectively by the dye molecules in the solution rather than by the adsorbed dye molecules at the electrode-solution interface, though only the excited adsorbed dye molecules can contribute to sensitisation²⁰.

To cope with this, a new type photocell has been developed (Fig. 1). The chemically modified SnO_2 or TiO_2 electrode with rhodamine B as a sensitizer is in contact with the transparent electrolyte solution containing a reducing agent as a supersensitizer. The procedure to modify the electrodes was similar to that previously described²¹, but synthesis of amide-linked rhodamine B by the dehydrative coupling of a carboxyl group of rhodamine B with propylamine modified electrodes was carried out using dicyclohexylcarbodiimide (DCC) as a dehydrating agent because it gives a better yield than the two-step reactions, involving the preparation of acid chloride of rhodamine B using thionyl chloride and the acylation of the propylamine



M = Sn or Ti; HOOC-RhB = rhodamine B

modified electrodes with the acid chloride. The propylamine modified TiO_2 electrodes were prepared in the same manner as the propylamine modified SnO_2 electrodes²¹ by the method developed by Murray *et al.*²². The surface reactions and confirmatory chemical tests were followed by X-ray photoelectron spectroscopy²¹.

Since the rhodamine B was covalently bound by silyl ether and amide bonds to the electrode surface, the dye remained almost permanently on the electrode surface, unless the pH value of the electrolyte solution was extremely low or high. Unwanted decomposition of the dye in the course of

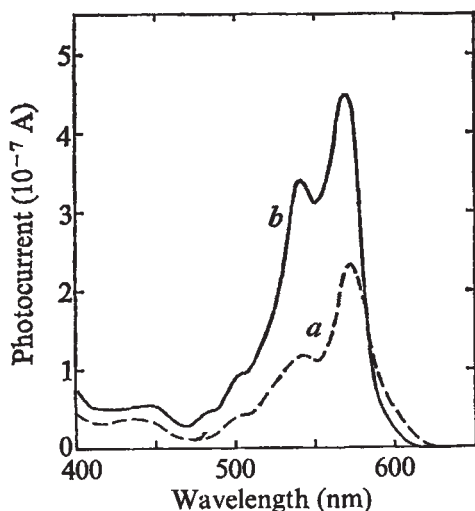


Fig. 2 Spectral dependence of the photocurrent for optically transparent SnO_2 electrodes polarised anodically at + 0.35 V against SCE, and illuminated through the electrode (Fig. 1). *a*, Photocurrent spectrum sensitised by a sub-monolayer of rhodamine B covalently bound to the electrode surface through amide bonding at the chemically modified SnO_2 electrode in contact with a transparent electrolyte (0.2 M Na_2SO_4) containing hydroquinone (5×10^{-3} M) as a supersensitizer. *b*, Photocurrent spectrum at the conventional unmodified SnO_2 electrode in contact with a rhodamine B solution (5×10^{-5} M rhodamine B + 5×10^{-3} M hydroquinone in 0.2 M Na_2SO_4).

the sensitisation process was almost entirely suppressed by the presence of the supersensitizer in the solution²³.

A typical sensitised photocurrent spectrum obtained with the modified electrode is shown in Fig. 2, as well as a spectrum obtained with the conventional untreated electrode in contact with rhodamine B solution containing the same amount of supersensitizer. The anodic photocurrents observed at the modified electrode, whose spectral dependence was quite close to that of the dye's absorption spectrum, is the same order of magnitude as those observed at the unmodified electrode, even though there is no dye in the solution. Since a monolayer coverage of the dye absorbs very little light^{24,25}, the modified electrodes are almost indistinguishable to the eye from the unmodified ones, and SnO_2 electrodes are transparent in the visible region. Accordingly, further improvements in efficiency can be easily attained by immersing many modified electrodes in the electrolyte solution in parallel containing colourless supersensitizers, and then by illuminating them perpendicularly.

Perhaps, the illumination of the electrode-electrolyte interface in a multiple internal reflection mode, (ATR, refs 24-27) will also be effective, since the chemically bound dyes on the surface can be excited by the 'evanescent wave', and loss of energy through reflections can be avoided in this mode. Extension of the useful spectral region may be possible by binding various sensitizers.

TETSUO OSA

MASAMICHI FUJIHARA

Pharmaceutical Institute,
Tohoku University,
Aobayama, Sendai 980, Japan

Received August 18; accepted October 9, 1976.

- 1 Fujishima, A., and Honda, K., *Nature*, **238**, 37-38 (1972).
- 2 Keeney, J., Weinstein, D. H., and Haas, G. M., *Nature*, **253**, 719-720 (1975).
- 3 Nozik, A. J., *Nature*, **257**, 383-385 (1975).
- 4 Frank, S. N., and Bard, A. J., *J. Am. chem. Soc.*, **97**, 7427-7433 (1975).
- 5 Fujishima, A., Kohayakawa, K., and Honda, K., *J. electrochem. Soc.*, **122**, 1487-1489 (1975).
- 6 Gissler, W., Lensi, P. L., and Pizzini, S., *J. appl. Electrochem.*, **6**, 9-13 (1976).
- 7 Wang, J. H., *Proc. natn. Acad. Sci. U.S.A.*, **62**, 653-660 (1969).
- 8 Tributsch, H., and Calvin, M., *Photochem. Photobiol.*, **14**, 95-112 (1971).
- 9 Tang, C. W., and Albrecht, A. C., *Nature*, **254**, 507-509 (1975).
- 10 Fong, F. K., and Winograd, N., *J. Am. chem. Soc.*, **98**, 2287-2289 (1976).
- 11 Gerischer, H., *J. electroanal. Chem.*, **58**, 263-274 (1975).
- 12 Archer, M. D., *J. appl. Electrochem.*, **5**, 17-38 (1975).
- 13 Wrighton, M. S., Morse, D. L., Ellis, A. B., Ginley, D. S., and Abrahamson, H. B., *J. Am. chem. Soc.*, **98**, 44-48 (1976).
- 14 Cronenmeyer, D. C., *Phys. Rev.*, **87**, 876-886 (1952).
- 15 Nagasawa, M., and Shionoya, S., *J. phys. Soc. Japan*, **30**, 1118-1123 (1971).
- 16 Fujishima, A., Hayashitani, E., and Honda, K., *Seisan Kenkyu*, **23**, 363-365 (1971).
- 17 Memming, R., *Photochem. Photobiol.*, **16**, 325-333 (1972).
- 18 Gerischer, H., Michel-Beyerle, M. E., Rebentrost, F., and Tributsch, H., *Electrochim. Acta*, **13**, 1509-1515 (1968).
- 19 Gerischer, H., and Tributsch, H., *Ber. Bunsenges. Phys. Chem.*, **72**, 437-445 (1968).
- 20 Fujishima, A., Watanabe, T., Tatsuoki, O., and Honda, K., *Chem. Lett.*, 13-18 (1975).
- 21 Fujihira, M., Matsue, T., and Osa, T., *Chem. Lett.*, 875-880 (1976).
- 22 Moses, P. R., Wier, L., and Murray, R. W., *Analyt. Chem.*, **47**, 1882-1886 (1975).
- 23 Watanabe, T., Fujishima, A., and Honda, K., *Ber. Bunsenges. Phys. Chem.*, **79**, 1213-1217 (1975).
- 24 Fujihira, M., Osa, T., and Hayano, S., *Rev. Polarography*, **21**, 160-161 (1975).
- 25 Mark, Jr., H. B., and Randall, E. N., *Symp. Faraday Soc.*, **4**, 157-172 (1970).
- 26 Hansen, W. N., *Analyt. Chem.*, **35**, 765-766 (1963).
- 27 Harrick, N. J., *Analyt. Chem.*, **36**, 188-191 (1964).

Cultural, language and geographical correlates of genetic variability in Andean highland Indians

THE genetic structure of primitive human populations has been studied for the purposes of understanding hominid evolution and developing more cogent genetic models. For this it is important not only to demonstrate genetic variability between populations, but also to explore the correlates of such variability. Using data on several Chilean Indian populations I have found that genetic differentiation can best be ascribed to geographical separations. Other correlates, such as culture and language, are of secondary importance when the relationship is studied simultaneously by a stepwise regression technique¹. Although I consider here a particular set of populations, the results obtained

Table 1 Correlation coefficients between genetic, geographical, cultural and linguistic variations among seven Chilean Indian tribes and their association with each tribe's extent of Indian ancestry

	Geographical distance	Cultural dissimilarity index	Linguistic dissimilarity index	Probability of Indian ancestry
Genetic	0.716*	0.775*	0.032	-0.018
Geographical	—	0.901*	0.323	-0.142
Cultural	—	—	0.147	0.016
Linguistic	—	—	—	-0.194

*Significant at 0.01 level.

seem to be fairly generally applicable². I have found an empirical relationship between geographical distance and gene identity for the Andean highland Indians discussed here.

There have already been attempts to relate genetic differentiation within a tribe or between several primitive tribes to linguistic, social, topographical and environmental components of the population²⁻⁴, but the conclusions were not always unequivocal. This is partly because all such correlates are themselves associated with one another, so that a simple analysis of variance as advocated by some researchers⁷, or a simple pairwise correlation⁴ may result in an unwarranted interpretation of the results.

The Chilean Indian tribes considered in this analysis are

organisation of the seven tribes. Another correlate was the linguistic variability which emerged from the nested hierarchical classification of Andean languages¹⁴. From the approximate geographical location of the tribes, the geographical distance between them was computed following standard physical transformation. As all the tribes have intermixed with Caucasoid invaders, the variation in the degree of Caucasoid admixture as determined from the ABO blood group were used to derive the probability of Indian ancestry for any two genes of the tribes.

To determine the nature of interrelations among these measures of variability, I first present the correlation coefficients between them using all the pairwise tribe-comparisons ($C_2=21$, see Table 1). Cultural and

Table 2 Results of a stepwise regression analysis identifying interrelationships between the different variability measures in Chilean Indian tribes

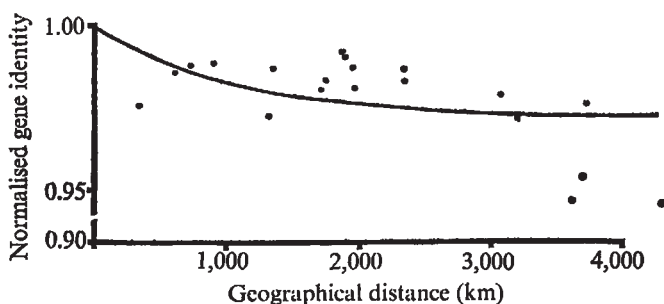
Variables entered	% Variance of genetic distance explained Cumulative	Marginal	F ratio (for each successive step)	Probability
Geography	51.31	51.31	20.02	$P < 0.001$
Culture	60.19	8.88	4.01	$0.05 < P < 0.1$
Language	61.48	1.29	0.57	$0.4 < P < 0.5$
Probability of Indian ancestry	61.52	0.04	0.02	$0.8 < P < 0.9$

the Aymara (17°S, 70°W), Atacameño (23°S, 67°W), Diaguita (22°30'S, 63°42'W), Pehuenche (38°S, 71°30'W), Mapuche (38°S, 72°30'W), Alacaluf (50°S, 75°W) and Yagan (55°36'S, 70°48'W), all so-called 'highland' Andean tribes. Gene frequencies for seven serological markers (ABO, MNSs, Rh, Diego, Duffy, Kell and haptoglobin) common to each population were extracted from the literature⁸⁻¹². To measure the genetic differentiation based on these markers, I computed the standard genetic distances between the different tribal groups, using the methods of Nei¹³.

As a first correlate of genetic differentiation, I constructed a cultural dissimilarity matrix taken directly from work³ involving an extensive literature survey of Andean culture. The basic features involved in cultural classification were the mode of subsistence, economics and social

geographical distances are almost linearly related ($r=0.901$). Linguistic dissimilarity was a poor correlate of genetic differentiation ($r=0.0320$, $P>0.85$). When geographical distance is used as the only regressor to explain genetic variabilities in the seven tribes, about 51.3% of the total variance is explained. All the other variables (cultural and linguistic variabilities, and variation in rates of admixture with Caucasians) explain only a marginal amount (10.2%). The result of such a regression analysis is shown in Table 2. This analysis also indicates that of all the variables, only geographical distance seems to be important, explaining the trend in genetic variabilities in these seven Andean populations. This agrees with theoretical predictions¹⁵. The relationship between genetic distance and geographical distance can be studied further by relating the probability of allelism between two homologous genes with the geographical distance by which two groups are separated. The probability of allelism, or the probability of gene identity between two populations, I_r , is given by $I_r = \exp(-D)$, where D is Nei's measure of standard genetic distance¹³. Figure 1 represents the empirical relationship between normalised gene identities (I_r) and geographical distances (r) for the 21 tribe comparisons. The least square fit of $I_r = 0.973 + 0.027 \exp(-r/1,000)$ is statistically satisfactory ($0.05 < P < 0.1$), although there are a few large deviations from expectations towards the tail of the curve.

There is also a high association between cultural and genetic variabilities in the seven tribes (Table 1), but this may be an artefact of the other two significant values in the same table. Such an association is also exhibited through the regression analysis, in the sense that in addition to geographical distance, cultural differences do not contribute significantly to the explanation of variability in genetic

Fig. 1 Empirical relationship between normalised gene identity and geographical distance (km) as studied from seven Chilean Indian tribes.

distances. The high correlation between culture and geography may also be interpreted from historical evidence, for within these populations geographical proximity was the only way of establishing cultural contact in the past. Furthermore, the material culture described here is not independent of ecological niche, especially because cultural distances are computed from information which includes subsistence.

A final observation is worth noting. Some researchers claim that a joint consideration of the linguistic and geographical divergencies at times indicates the presence of selective pressures, for example, linguistically distant but geographically close groups showing a great deal of genetic identity may be an indicator of selection as an evolutionary force³. In view of my analysis, such statements seem to be premature, for linguistic divergence can be the result of only recent historic developments of such tribal groups, whereas genetic differentiation by gene frequency changes is known to be a very slow process. Furthermore, in close proximity, exchange of a few mates between populations can never be ruled out. Theoretical predictions reveal that such migration need not be great to bring two neighbouring populations genetically close together⁴.

This research was supported by grants from the USPHS.

RANAJIT CHAKRABORTY

Center for Demographic and Population Genetics,
University of Texas Health Science Center,
Houston, Texas 77030

Received June 18; accepted September 27, 1976.

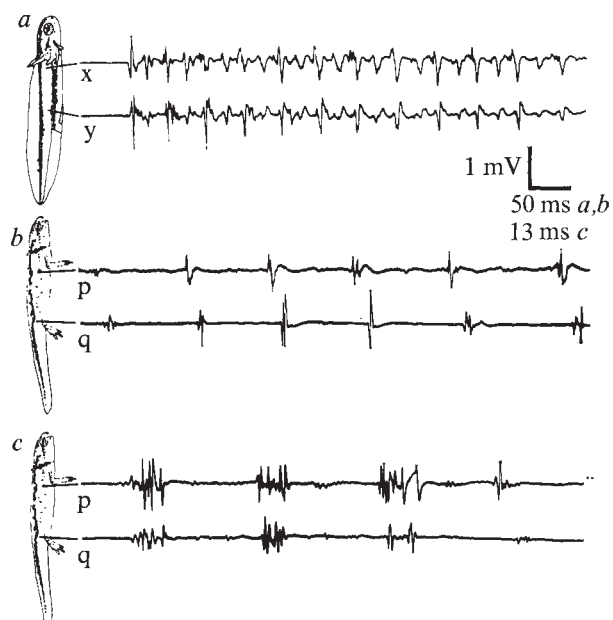
- 1 Snedecor, G. W., and Cochran, W. G., *Statistical Methods* (Iowa State University Press, Ames, 1967).
- 2 Roychoudhury, A. K., *Am. J. phys. Anthropol.*, **42**, 449-454 (1975).
- 3 Balakrishnan, V., Sanghvi, L. D., and Kirk, R. L., *Genetic Diversity among Australian Aborigines* (Australian Institute of Aboriginal Studies, Canberra, 1975).
- 4 White, N. G., and Parsons, P. A., *Nature*, **261**, 223-224 (1976).
- 5 Chakraborty, R., Blanco, R., Rothhammer, F., and Llop, E., *Soc. Biol.*, **23**, 73-82 (1976).
- 6 Rhoads, J. G., and Friedlaender, J. S., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 2247-2250 (1975).
- 7 Friedlaender, J. S., *Patterns of Human Variation: The Demography, Genetics, and Phenetics of Bougainville Islanders* (Harvard University Press, Cambridge, 1975).
- 8 Etcheverry, R., et al., *Rev. Med. Chile*, **95**, 599-604 (1967).
- 9 Etcheverry, R., et al., *Rev. Med. Chile*, **95**, 605-608 (1967).
- 10 Matson, G. A., Sutton, H. E., Swanson, J., and Robinson, A., *Am. J. phys. Anthropol.*, **25**, 13-33 (1966).
- 11 Matson, G. A., Sutton, H. E., Etcheverry, R., Swanson, J., and Robinson, A., *Am. J. phys. Anthropol.*, **27**, 157-194 (1967).
- 12 Matson, G. A., Sutton, H. E., Swanson, J., and Robinson, A., *Am. J. phys. Anthropol.*, **30**, 61-84 (1969).
- 13 Nei, M., *Am. Nat.*, **106**, 283-292 (1972).
- 14 Greenberg, J. H., in *Selected Papers of 5th International Congress of Anthropological and Ethnological Sciences* (edit. by Wallace, A. F. C.), (University of Pennsylvania Press, Philadelphia, 1960).
- 15 Maruyama, T., in *Genetic Structure of Populations* (edit. by Morton, N. E.), 80-81 (University of Hawaii Press, Honolulu, 1973).
- 16 Chakraborty, R., and Nei, M., *Theor. Pop. Biol.*, **5**, 460-469 (1974).

dogfish^{6,7} has not been related to this mechanical problem. I present here a preliminary report of electromyographic studies of developmental stages of the palmate newt, *Triturus helveticus* (Razoumowsky), and adult tench, *Tinca tinca* L., which show that undulatory propulsion in some animals does not involve waves of contraction and that the "waves" recorded in others function in reducing lateral oscillation anteriorly and adapting tail flexibility to different swimming speeds.

From the time when swimming by waves of bending is first possible in the newt embryo until the time of normal hatching there is no longitudinal delay in muscle activity along the body during the tailbeat. The electromyograms suggest that only the anterior myotomes are used (not those in the tail nor, at first, those in the lower trunk) and these contract together on either side (Fig. 1a). Recent ultrastructural studies have suggested that there is some electrical couplings between myotomal muscle fibres in amphibian embryos which persists for a short time after the development of swimming⁸⁻¹¹. Such coupling may be involved in the synchronisation of contraction along the trunk. The boundary between myotomes which seem to be active and those which seem to be inactive in swimming moves slowly back to the junction of trunk and tail by the end of embryonic life. But the inactivity of the more posterior myotomes survives the development of the capability for involvement in struggling movements, suggesting that it is a functional rather than merely a developmental feature.

The mechanics of this form of locomotion are easy to understand in outline from the simple oscillation of a flexible plate under water. Bending moments can be genera-

Fig. 1 Electromyograms of swimming activity in *T. helveticus*, recorded with 25- μ m diameter wire electrodes (tip position indicated on the left). a, (Monopolar) from a stage-37 embryo, 7.4 mm long; b and c, (bipolar) from a stage-55c larva, 29 mm long (see ref. 8 for developmental stages). In the thinner embryo the contralateral muscle activity is recorded at about half the amplitude of that from the side in which the electrodes are implanted. In the larva the contralateral activity is of much lower amplitude in the recordings. In (a) each of the large amplitude, spike-like events corresponds to a beat of the tail to the right, each of the lower amplitude intervening events to a beat to the left. There seems to be no delay in the occurrence of muscle activity down the length of the body comparable with that in (b) where there is a clear delay in the arrival of activity at electrode q after its cessation at electrode p. c, Violent swimming activity in the same animal, with a return to simultaneity except in the last, low amplitude beat. (The vertical gain on the trace from electrode p is five times that on the other traces which correspond to the scale-bar.)



Undulatory swimming with and without waves of contraction

IT HAS been accepted generally that the waves of bending observed in swimming vertebrates are produced by contraction on the concave side of each bend of the body at any point within the cycle of movement¹⁻³. "Waves of contraction" are imagined to pass down the serial myotomes with the velocity of the propagated bends. Another concept, designed primarily to explain snake locomotion⁴, differs by a 90° shift in the timing of contraction on the bending wave, accounting for propulsion but resulting in illogical analysis of fish movement⁵. The former interpretation gives no explanation of the source of locomotory thrust, assuming falsely that the muscular activity required for a given body position is similar whether the position is held or passed through in locomotion. The dominant responsibility of locomotory activity is the propagation rather than the formation of bends, but how is propagation achieved against external resistance? Relevant electromyography of spinal

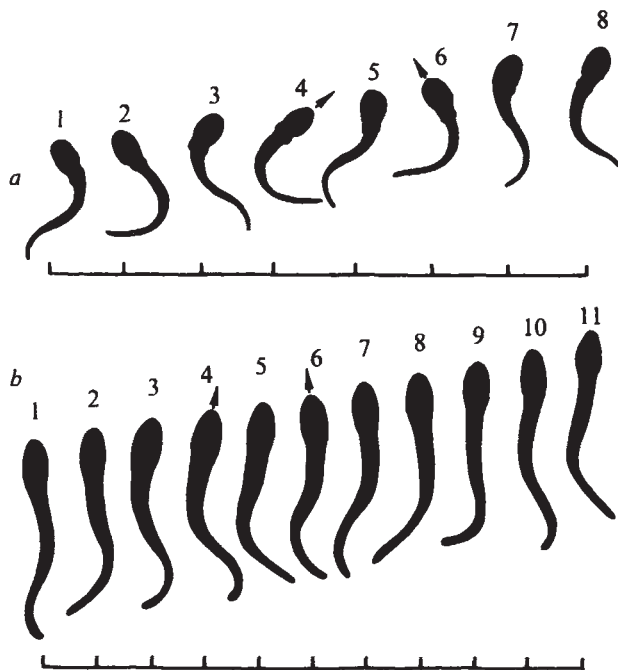


Fig. 2 Two series of tracings from ciné film of the swimming of *T. helveticus*: *a*, from a stage-37 embryo; *b*, from a stage-55c larva. Both sequences were filmed at 64 frames per s and the frame numbers for each sequence are given. Separation of each position from the baseline represents real movement in space. Each position has been displaced to the right from the preceding one by a fixed distance perpendicular to the baseline (shown by the scale divisions). The reduction of the oscillation of the head axis in (*b*) is indicated by the arrows showing the extreme positions at frames 4 and 6. The limbs have not been drawn in (*b*). The lengths of the animals are given in Fig. 1.

ted anteriorly and conducted along the tail by passive elastic tension. The form of the propagated waves will then depend on the animal's structural characteristics—particularly its flexibility and lateral hydrodynamic resistance—and their longitudinal distribution. The presence of a fin around the tail facilitates the development of both high flexibility and high lateral resistance in that region. These characteristics, combined with muscular passivity, can be seen to be responsible for the whip-like behaviour of the oscillated tail. In this early form of swimming the longitudinal axis of the head may swing through as much as 120° in the horizontal plane within the swimming cycle. During embryonic life the rotation of the head axis is reduced slightly (Fig. 2*a*) but is still measured in tens of degrees.

Early in larval life the myotomal contractions begin to show longitudinal delay, the propagation of some kind of "wave" (Fig. 1*b*). The characteristics of this wave are widely variable, as is the swimming behaviour of the larva, and in the most violent swimming the wave phenomenon is replaced by simultaneity of contraction along the myotomes (Fig. 1*c*). Associated with the wave of contraction is the reduction of the oscillation of the axis of the head to only a few degrees (Figs 2*b* and 4). In the newt larva, violent, "longitudinally simultaneous" swimming is accompanied by a return to wide lateral oscillation of the head.

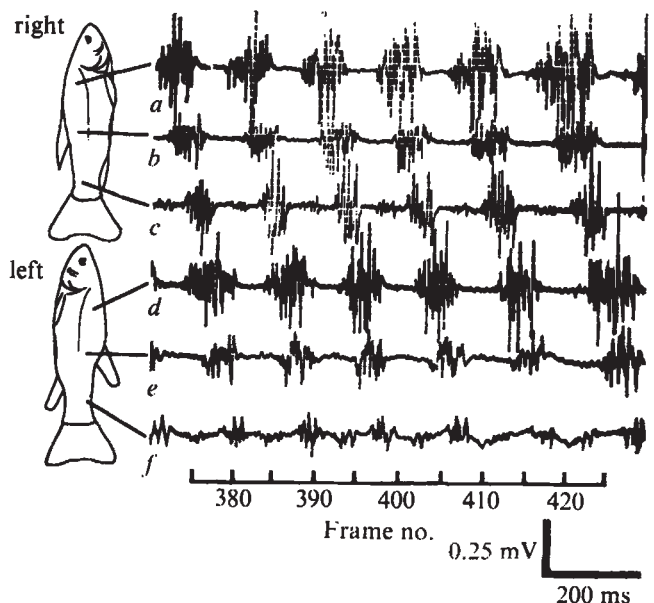
The wave phenomenon is more easily studied in a larger animal. The electromyogram of the tench in Fig. 3 shows that the delay in the initiation of muscle activity down the body is about twice that of the cessation. This means that the 'burst' of activity becomes shorter, moving caudally along the animal. There is also a period within each half-cycle of movement in which muscle activity is recorded simultaneously over the major part of the myotomal column. If the electromyogram is compared with the movements of the

animal (Fig. 4), periods of simultaneous lateral activity correspond with what may be called the S position of the body, the point at which conventional theory would least predict it. The contralateral overlap of activity occurs in those parts of the cycle in which the fish is approaching the C shape which is fully realised in the swimming of the newt embryo (Fig. 2*a*), but not quite achieved in the swimming of the tench (or the gentle swimming of the larval newt—Fig. 2*b*).

So although waves of contraction are not necessary merely for the propagation of waves of bending, they are recorded in the adult swimmer and they require an alternative explanation. The early initiation of contraction on one side of the anterior trunk, while the tail is still moving towards the other side, prevents the full lateral deflection of the head (Fig. 4). A similar effect is achieved by the newt larva in Fig. 1*b* with early relaxation anteriorly without contralateral contraction. This reduction of anterior movement may have a hydrodynamic significance¹², and certainly will have a behavioural significance in stabilising the cephalic sensory reception during locomotion, particularly in animals which hunt active prey.

In combination with this anterior function of the wave is the more complex concern with providing for various frequencies of tailbeat. The passive tail of the embryo would only function properly over a narrow range of frequencies suited to its mechanical characteristics, and this seems to be all that it is required to do. The embryo swims only at relatively high frequencies (usually 15–25 Hz). The activity of the tail muscle of the more developed animal has to alter the mechanical properties of the tail for wave conduction at different frequencies, for different frequencies entail different lateral velocities and accelerations, which in turn require different bending moments to be generated in the tail if the effective waveform is to be maintained. To do this economically, the mechanical properties of the tail must be suited for passive conduction of waves at the lowest frequencies required (perhaps 1 or 2 Hz) and then be 'stiffened'

Fig. 3 A 1-s portion of electromyogram from a free swimming tench during moderately fast cruising, recorded with six bipolar wire electrodes. The four forward electrodes were close under the skin, those in the tail peduncle (*c* and *f*) were deeper (so that they show something of the contralateral activity). The electrode tip positions are indicated on the left. The frame numbers of the horizontal axis refer to the ciné film which is partly traced out in Fig. 4. There is a delay of about 1.0 s m^{-1} in the initiation of muscle activity down the body, which is about twice the delay in its cessation (about 0.5 s m^{-1}). The overall length of the fish was 16.5 cm.



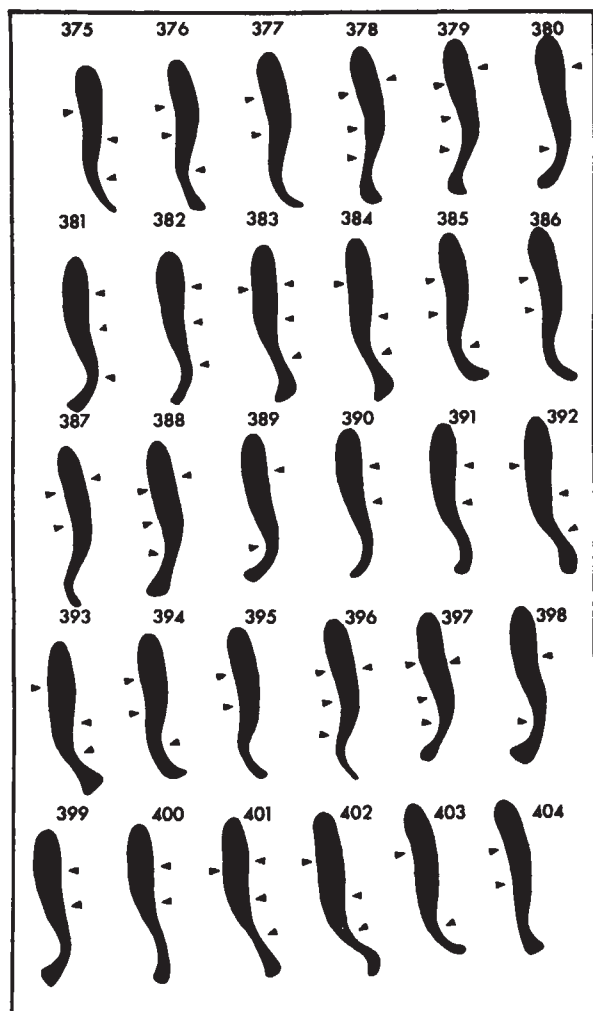


Fig. 4 A series of tracings from cine film of the swimming of the adult tench electromyographically recorded in Fig. 3. The frame numbers above each position correspond to the scale given in Fig. 3. The arrowheads indicate the approximate sites of electrodes which were showing electrical activity in the muscle at the time of exposure of the film. (The form of the fish is at times slightly distorted by disturbance of the water surface.) The S position of the swimming cycle is characterised by activity only on one side or all down one side. The occurrence of contraction at the anterior of one side and the posterior of the other is associated with the approach of the tip of the caudal fin to its extreme lateral displacement. This condition is often preceded by activity recorded anteriorly appearing on both sides simultaneously. Camera speed was approximately 64 frames s^{-1} .

by muscular contraction during faster swimming movements. In this way the level of muscle activity is directly proportional to speed, rather than inversely, as would be the case if the tail were mechanically suited to the highest speed and actively flexed for lower frequencies. When these adaptations are incorporated, the wave of bending is produced partly passively, partly actively, and the role of the wave of contraction becomes complicated by involvement in modification of the propagation of the bend for the subtle control of movement.

This work was carried out in the Department of Zoology, University of Bristol, and supported by the SRC. I thank Drs Q. Bone, A. Roberts, and H. P. Whiting for discussion.

ANDREW R. BLIGHT

Department of Physiology and Biophysics,
New York University Medical Center,
550 First Avenue,
New York, New York 10016

Received July 7; accepted September 13, 1976.

¹ Coghill, G. E., *J. comp. Neurol.*, **19**, 83-105 (1909).

- ² Breder, C. M., *Zoologica*, New York, **4**, 159-297 (1926).
- ³ Gray, J., *J. exp. Biol.*, **10**, 88-104 (1933).
- ⁴ Gray, J., *Q. Jl microsc. Sci.*, **94**, 551-578 (1953).
- ⁵ Gray, J., in *Essays in Marine Biology* (edit. by Marshall, S. M., and Orr, A. P.), 1-16 (Oliver and Boyd, Edinburgh, 1953).
- ⁶ Grillner, S., *Expl Brain Res.*, **20**, 459-470 (1974).
- ⁷ Grillner, S., Perret, C., and Zangger, P., *Brain Res.*, **109**, 255-269 (1976).
- ⁸ Gallien, L., and Bidaud, O., *Bull. Soc. zool. Fr.*, **84**, 22-32 (1959).
- ⁹ Keeter, J. S., Pappas, G. D., and Model, P. G., *Devl Biol.*, **45**, 21-33 (1975).
- ¹⁰ Hayes, B. P., *Anat. Embryol.*, **147**, 345-354 (1975).
- ¹¹ Blackshaw, S. E., and Warner, A. E., *J. Physiol., Lond.*, **246**, 73P-74P (1975); *ibid.*, **255**, 209-230 (1976).
- ¹² Lighthill, M. J., *J. Fluid Mech.*, **9**, 305-317 (1960).

Effect of caffeine on coffee drinking

CAFFEINE is one of the world's most widely-used drugs and coffee drinking is a major form of its use¹. The levels of caffeine intake in coffee drinking have significant, predominantly stimulatory, physiological and behavioural effects². Surprisingly, the role of caffeine in the self-administration of coffee has not yet been studied. This research indicates that coffee drinkers drink more low-caffeine coffee than non-decaffeinated coffee.

In study 1, Instant Maxwell House Coffee (3.44% caffeine) and Instant Sanka Brand 97% Caffeine Free Coffee (0.14% caffeine) were mixed to produce doses of 100, 50, and 25 mg per cup of coffee. The coffees were packaged in individual-serving foil pouches of 2.94 gm. To avoid severe withdrawal symptoms (headache, fatigue or irritability)^{3,4} and to provide the possibility of maintaining normal caffeine consumption, smaller caffeine doses were avoided. (About 33 cups of decaffeinated coffee are needed to give 100 mg caffeine—the non-decaffeinated dose.)

Subjects were told that they would be receiving different types of coffee without additives. The study took 3 d for each subject—a day on each caffeine dose. The order of presentation was balanced. Daily, subjects received a generous, counted supply of coffee packets and a short questionnaire to assess how they were feeling and how the coffee tasted (good-bad, weak-strong, bitter-sweet, on four-point bipolar scales). Unused and used packets were to be returned. A count of the packets gave a check on reports of coffee intake. Consumption of other coffee, any tea, cola drinks, cocoa, chocolate and certain headache preparations was forbidden: all contain caffeine.

The volunteer subjects (seven females, five males) had an average age of 41 (range: 27-66) and by pre-experimental reports drank 4.5 cups of coffee per day (range: 3-6.5).

Table 1 shows a tendency for coffee intake to increase as caffeine content decreases. Subjects did not 'taste' the caffeine manipulation ($P > 0.5$). Changes in an index of how well the subjects were feeling (by self-report) correlated with changes in coffee consumption. The increased intake

Table 1 Average coffee intake (cups per person) as a function of the caffeine content of the coffee in three studies ($\pm$ s.d.)

Study	Caffeine content (mg)		
	100	50	25
1	4.1*	5.0*	4.5
	± 1.2	± 1.7	± 1.1
2	1.2†‡	2.0†	2.1‡
	± 0.43	± 0.06	± 0.06
3	2.6§	3.0§	3.1
	± 0.09	± 0.19	± 0.15

Means with same superscripts are different by two-tailed t tests, $P < 0.05$. In study 1, t tests for correlated samples were used; mean combined 50- and 25-mg intake was greater than 100-mg intake, $P < 0.05$; the unexpected quadratic trend, $F(1,22) = 6.10$, $P < 0.05$, may arise from debility from headaches in three subjects on 25-mg caffeine; if they are excluded, the trend disappears and 25- and 50-mg coffee intake becomes almost identical. In studies 2 and 3, t tests for independent samples are used for a rough estimate of statistical significance⁵.

of 50-mg relative to 100-mg coffee was associated with increased well-being ($r(10) = +0.70$, $P < 0.02$); the decreased intake of 25-mg relative to 50-mg coffee was associated with decreased well-being ($r(10) = +0.67$, $P < 0.05$).

Studies 2 and 3 treated an institutional coffee pot as the 'subject' in a single subject experiment. If the drinkers comprise constant groups, then the rate of emptying the pots—as a function of caffeine doses—is a valid measure of caffeine's effect on coffee intake.

Study 2 lasted from 0800 to 1100—the 'breakfast pot'—on Wednesdays and Thursdays for 4 weeks on a 40-cup coffee pot located in the dining room of a university hall of residence. There were about 12 regular coffee drinkers (six males, six females—all undergraduates). Roast and ground Maxwell House Coffee and Sanka Brand 97% Caffeine Free Coffee were mixed to yield 100-, 50-, and 25-mg caffeine doses per cup of coffee (week 1=100, week 2=50, week 3=25, and week 4=100 mg). The coffee left in the pot was measured with a ruler. Table 1 shows the estimated average coffee intake per person (each person weighted equally).

Study 3 used the 70-cup coffee maker used by ~ 25 secondary school teachers (~ 66% are female). The study lasted 3 weeks (week 1=100, week 2=25, and week 3=50 mg), from 0800 to 1200 on Tuesdays, Wednesdays, and Thursdays. Subjects were not asked to change their usual drinking habits: many had coffee before coming to work. Table 1 shows the results (adjusted for teacher absenteeism).

These studies were combined by transforming the results to standardised (or z) scores, see Fig. 1. Caffeine is shown to be a determinant of the frequency of coffee consumption. Though the regulation of caffeine intake is far from perfect (total caffeine intake decreases), it is impressive, on considering a few factors. Habits of intake may be slow to

change. Coffee drinking is also water drinking: water regulatory systems could prevent increases in fluid intake. Perhaps non-decaffeinated coffee provides surpluses of caffeine, such that 50-mg coffee ('half-caffeine' coffee) with only small increases in consumption is able to satisfy actual caffeine needs⁶. Too small doses of stimulants have been known to not sustain self-administration⁷: this may explain the failure of 25-mg coffee to be consumed more than 50-mg coffee. A half-caffeine mixture might be recommended to those who are trying to cut down on caffeine intake, but who still desire some of the benefits of caffeine.

I thank the General Foods Corporation, Technical Center, Tarrytown, N.Y., and K. R. Hirsh for supplies and technical assistance; and D. Ebb, S. Doyle, L. McLane, S. Verostyk, K. MacHan, and Coginchau Regional High School, Durham, Connecticut.

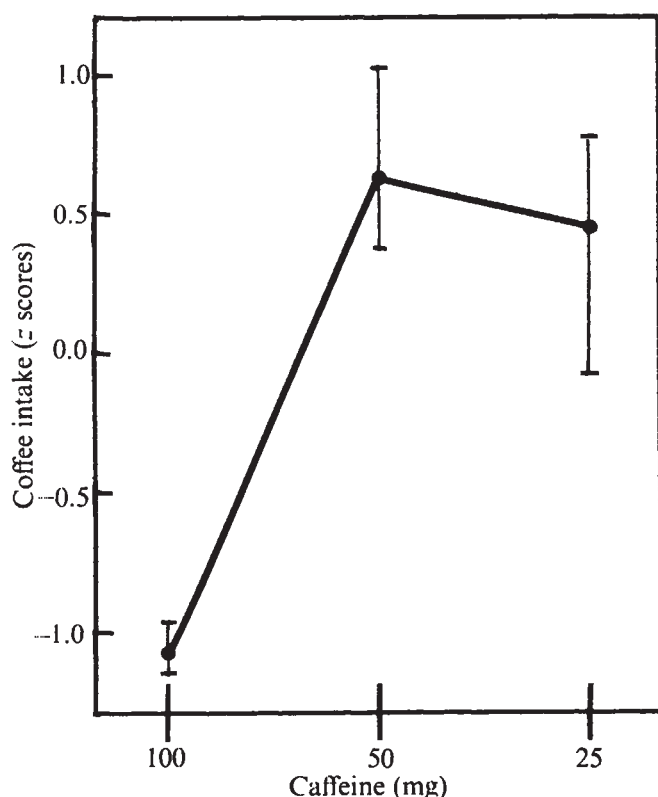
LYNN T. KOZLOWSKI

Department of Psychology,
Wesleyan University,
Middletown, Connecticut 06457

Received July 30; accepted September 27, 1976.

- 1 Brecher, E. M., *Licit and Illicit Drugs* (Little, Brown, Boston, 1972).
- 2 Ritchie, J. M., in *The Pharmacological Basis of Therapeutics* (edit. by Goodman, L. S., and Gilman, A.), 358–370 (Macmillan, New York, 1970).
- 3 Dreisbach, R. H., and Pfeiffer, C., *J. lab. clin. Med.*, **28**, 1212–1219 (1943).
- 4 Goldstein, A., Kaiser, S., and Whitby, O., *Clin. pharmacol. Ther.*, **10**, 488–497 (1969).
- 5 Johnson, L. C., and Lubin, A., in *Handbook of Psychophysiology* (edit. by Greenfield, N. S., and Sternbach, R. A.), 125–158 (Holt, Rinehart, Winston, New York, 1972).
- 6 Kozlowski, L. T., *Psychopharmacology (Berl.)*, **47**, 165–168 (1976).
- 7 Deneau, G. A., and Inoki, R., *Ann. N.Y. Acad. Sci.*, **142**, 277–279 (1967).

Fig. 1 Standardised coffee intake scores (mean = 0, s.d. = 1) on average number of cups of coffee drunk per person as a function of caffeine content of coffee (log scale). Each study was standardised separately, and the range is indicated. The caffeine manipulation influenced coffee intake ($F(2,4) = 16.01$, $P < 0.02$). 50-mg coffee intake was greater than 100-mg coffee intake, $P < 0.02$; 25-mg intake was greater than 100-mg intake, $P < 0.05$; 50- and 25-mg coffee intake did not differ.



Vaccination against bovine tropical theileriasis (*Theileria annulata*)

SINCE the reports of Tsur¹ and Brocklesby and Hawking² on the multiplication of *Theileria annulata* schizonts in spleen explant cultures in plasma clot, the schizonts have been grown successfully for months in tissue culture monolayers^{3–5} and suspension cultures⁶. Successive cultivation of schizonts in tissue culture has been shown to bring about attenuation of the parasite^{7–10}, and suspensions of attenuated schizonts have been used as vaccine to immunise cattle against *T. annulata* in Israel⁶ and Iran¹¹. Reports^{8,11} indicate that the protective effect of the vaccine has been tested against challenge with schizonts but not against the natural tick-induced infection. As there may be some difference in the antigens of schizont and sporozoite (the injected stage of the parasite in the tick), we have evaluated the immunity produced by the vaccine against the infection caused by tick infestation.

Schizont-infected lymphoid cells were obtained by biopsy of a lymph gland of a calf infested with the tick *Hyalomma dromedarii* and infected with virulent Hissar strain¹² *T. annulata*. The cells were cultured in Eagle's MEM (Hanks' base), supplemented with 10% calf serum. Subcultures were made every second or third day and vaccine was prepared from the fifteenth passage.

Crossbred male calves reared in tick-proof rooms were vaccinated when two months old. Fifteen calves were injected subcutaneously at the base of the ear near the parotid lymph node with 1×10^6 schizonts of the fifteenth subpassage. After 35 d, they were challenged with the homologous strain of *T. annulata* by infestation with 10 infective ticks (*H. dromedarii*). The ticks were released on the ear, which was enclosed in a cloth bag. Eight similarly raised susceptible control calves were also infected.

The rectal temperature of the calves was recorded daily. Giemsa-stained smears of blood were examined daily, and of the biopsy material from the regional parotid lymph node

Table 1 Course of infection induced by inoculation of infected lymphoid cells

Calf no.	First onset of fever (d)	Duration of fever (d)	Maximum temperature (°C)	Commencement of swelling of the regional parotid lymph node (d)	% Schizont-infected cells in the lymph node biopsy material	First appearance of intraerythrocytic piroplasms (d)	Maximum parasitaemia
6	17	2	39.5	No swelling	Nil	—	Nil
7	No fever	—	—	No swelling	Nil	—	Nil
8	No fever	—	—	No swelling	Nil	—	Nil
9	No fever	—	—	No swelling	Nil	—	Nil
10	No fever	—	—	No swelling	Nil	—	Nil
200	No fever	—	—	No swelling	Nil	29	<1
191	13	5	40.4	No swelling	Nil	24	<1
192	15	1	39.4	No swelling	Nil	29	<1
193	15	4	40.0	No swelling	Nil	17	<1
197	16	4	40.0	No swelling	Nil	—	—
194	13	5	39.7	15	Nil	17	<1
195	15	4	40.4	15	<1	16	<1
196	13	6	40.4	15	Nil	20	<1
198	16	1	40.0	15	Nil	20	<1
199	15	2	40.1	15	Nil	—	—

Table 2 Course of *T. annulata* infection induced by infestation of unvaccinated control calves

Calf No.	First onset of fever (d)	Duration of fever (d)	Maximum temperature (°C)	Commencement of swelling of the regional parotid lymph node (d)	% Schizont-infected cells in lymph node biopsy material	First appearance of intra-erythrocytic piroplasms (d)	Maximum parasitaemia (%)	Results
11	10	9	41.1	—	14	12	15	Survived
12	11	10	40.9	—	8	11	20	Died on day 25
13	11	10	41.0	—	12	12	25	Died on day 22
14	9	11	41.0	—	12	10	12	Survived
232	6	Until death	41.3	8	2	13	60	Died on day 15
234	9	Until death	41.1	8	25	13	80	Died on day 15
235	9	Until death	41.1	7	10	13	80	Died on day 15
237	9	Until death	41.1	7	10	13	80	Died on day 16

on alternate days beginning a week after the vaccination/challenge. Haemoglobin (Hb) content and packed cell volume (PCV) were determined twice a week.

The reactions to vaccination are summarised in Table 1. Ten of fifteen calves developed fever which commenced between days 13 and 16 and lasted for 1–6 d. The maximum temperature was 40.4 °C. Transient swelling of a regional parotid lymph node was observed in five calves. Parasitised lymphocytes (less than 1%) in the regional lymph gland appeared only in one calf (No. 195) on day 17. Intraerythrocytic piroplasms were observed in 8 of 15 calves between days 17 and 29 for the first time, and later at irregular intervals. Parasitaemia always remained less than 1%. Hb content and PCV of all calves remained normal, and all the calves survived.

When challenged, immunised calves showed no clinical or parasitological reaction, except for one calf (No. 199) which showed less than 1% parasitaemia on day 16. Hb content and PCV remained normal and the blood and regional lymph gland remained free of the parasites. But all eight calves of the control group suffered from the acute type of the disease and six of them died between days 15 and 25 (Table 2). The disease was characterised by high fever which started in 6–11 d, and lasted for 9 d or until death. The intraerythrocytic form of the parasite appeared in 10–13 d. Peak parasitaemias varied from 12 to 80%. The regional lymph gland was swollen in four of eight calves, although infected lymphoid cells (2–25%) were seen in the gland of all the calves.

We conclude therefore that all 15 vaccinated calves became fully immune to the severe tick challenge which killed six of eight unimmunised calves.

The vaccine produced no reaction in five calves and very mild clinical and/or parasitological reaction in the remaining 10, as also observed by Pipano and Israel⁹ with their

10th–15th passage material which at the 50th–120th passage produced no reaction. We are continuing to subculture the parasite with a view to further attenuation.

We thank the Indian Council of Agricultural Research for financial support.

B. S. GILL
Y. BHATTACHARYULU
D. KAUR
AVTAR SINGH

Punjab Agricultural University,
Ludhiana 141004, India

Received June 16; accepted August 17, 1976.

- 1 Tsur, I., *Nature*, **156**, 391 (1945).
- 2 Brocklesby, D. W., and Hawking, F., *Trans. R. Soc. trop. Med. Hyg.*, **52**, 414 (1958).
- 3 Tsur, I., and Adler, S., *Refuah Vet.*, **19**, 225 (1962).
- 4 Hulliger, L., *J. Protozool.*, **12**, 649 (1965).
- 5 Hulliger, L., Wilde, J. K. H., Brown, C. G. D., and Turner, L., *Nature*, **203**, 728 (1964).
- 6 Hooshmand-Rad, P., and Hashemi-Fesharki, R., *Arch. Inst. Razi*, **20**, 85 (1968).
- 7 Tsur, I., and Pipano, E., *J. Protozool.*, **13**, Suppl. 33 (1966).
- 8 Pipano, E., and Tsur, I., *Refuah Vet.*, **23**, 133 (1966).
- 9 Pipano, E., and Israel, V., *J. Protozool.*, **18**, Suppl. 37 (1971).
- 10 Pipano, E., Klopfer, U., and Cohen, R., *Res. vet. Sci.*, **15**, 388 (1973).
- 11 Hashemi-Fesharki, R., and Shah-Del, F., *Am. J. vet. Res.*, **34**, 1465 (1973).
- 12 Gill, B. S., Bhattacharyulu, Y., and Kaur, D., *Res. vet. Sci.*, **21**, 146 (1976).

Non-adaptive rejection of small tumour inocula as a model of immune surveillance

THOMAS's thesis¹ that it is a "universal requirement of multi-cellular organisms to preserve uniformity of cell type", led him to suggest that the phenomenon of homograft rejection represented the primary mechanism for natural defence against neoplasia. Burnet² viewed this surveillance of lethal oncogenic mutations as a function of the thymus-dependent adaptive immune system. Prehn^{3,4} and others⁵⁻⁷ have criticised the Burnet model of immune surveillance because it is unable to account for several important observations: (1) the adaptive

Table 1 Mice rejecting small syngeneic inocula in a primary challenge fail to exhibit memory on secondary challenge

Experiment	Tumour	Route of administration§	Inoculum size	%Cloning efficiency¶ ±s.e.m.	Tumour incidence on primary challenge		Large inoculum second challenge**		Small inoculum second challenge††	
					Number	%	Latency† (d±s.e.m.)	Survival (d±s.e.m.)	Number	%
1.	L5178Y*	Subcutaneous	5 × 10 ¹	47 ± 5	4/15	26	—	—	—	—
2.	L5178Y*	Subcutaneous	5 × 10 ¹	25 ± 7	8/40	20	12.5 ± 2.0	35.4 ± 1.0	—	—
		Intraperitoneal	5 × 10 ¹	25 ± 7	15/15	100	—	—	—	—
		Control	0	—	0/15	0	13.5 ± 2.5	36.4 ± 2.2	—	—
3.	P-815-X2*	Subcutaneous	5 × 10 ¹	44 ± 2	11/40	27.5	8.0 ± 1.0	33.8 ± 1.3	—	—
		Intraperitoneal	5 × 10 ¹	44 ± 2	15/15	100	—	—	—	—
		Control	0	—	0/15	0	7.9 ± 1.0	32.1 ± 2.3	—	—
4.	L5178Y†	Subcutaneous	10 ¹	—	1/8	12.5	—	33 ± 2	—	—
		Subcutaneous	5 × 10 ¹	—	2/8	25	—	40 ± 3	—	—
		Subcutaneous	10 ²	—	3/8	37.5	—	33 ± 5	—	—
		Control	0	—	0/5	0	—	45 ± 2	—	—
		Intraperitoneal	10 ¹	—	13/15	87	—	—	—	—
5.	P-815-X2†	Subcutaneous	10 ¹	—	4/8	50	—	38 ± 6	—	—
		Subcutaneous	5 × 10 ¹	—	2/8	25	—	39 ± 3	—	—
		Subcutaneous	10 ²	—	8/8	100	—	—	—	—
		Control	0	—	0/5	0	—	41 ± 5	—	—
		Intraperitoneal	10 ¹	—	8/8	100	—	—	—	—
6.	SL2†	Subcutaneous	10 ¹	—	3/8	37.5	—	37 ± 6	—	—
		Subcutaneous	5 × 10 ¹	—	6/8	75	—	32 ± 5	—	—
		Subcutaneous	10 ²	—	4/8	50	—	30 ± 7	—	—
		Control	0	—	0/5	0	—	35 ± 3	—	—
7.	1509a*	Subcutaneous	10 ²	—	3/20	15	—	—	3/14	21
		Subcutaneous	10 ³	—	5/20	25	—	—	2/10	20
		Subcutaneous	10 ⁴	—	13/20	65	—	—	3/6	50
		Subcutaneous	10 ⁵	—	20/20	100	—	—	—	—
		Intraperitoneal	10 ²	—	5/5	100	—	—	—	—
		Intraperitoneal	10 ³	—	5/5	100	—	—	—	—
		Intravenous	10 ²	—	0/10	0	—	—	—	—

*Tumours obtained from *in vitro* cultures for both primary and secondary challenge.

†Tumours obtained from ascitic passage for both primary and secondary challenge.

‡Appearance of palpable tumour after subcutaneous inoculation.

§All inoculations were 0.1 ml of washed cells. Subcutaneous inoculations were made in mid-low back after removing the fur.

¶After thorough washing, the tumour was resuspended in Fischer's medium containing 15% foetal calf serum. Aliquots of 2.0 ml containing 5 × 10¹ cells were added to 3 ml of 2% Noble special agar at 44 °C. After thorough mixing, the suspension was placed in a 37 °C incubator at 5% CO₂ in a humid atmosphere. Visible clones were counted after 2 weeks in culture. Cloning efficiency was calculated from a minimum of 10 replicates.**An aliquot of 10⁵ tumour cells in 0.1 ml Fischer's medium was inoculated subcutaneously in the same site as the primary challenge.

††Tumour frequency was assessed 25 d after subcutaneous challenge with the same number of tumour cells as the primary inoculum.

immune system, with only rare exceptions, is unable to eliminate incipient or induced tumours rather than functioning as a surveillance mechanism^{4,6-8}; (2) the immune response to tumours is as likely to result in immunosuppression and enhancement of tumour growth as the destruction of those tumours^{6,8,9}; and (3) the incidence of spontaneous tumours in athymic mice is insignificant⁵. To explain these data one must either assume that surveillance does not exist, or one must postulate that surveillance exists in a form other than the thymus-dependent adaptive immune response. We have attempted to examine the hypothesis that surveillance exists and is a non-thymus-dependent phenomenon by an experimental model in which the fate of a small tumour inoculum, simulating an oncogenic mutation, is examined after manipulation of the host's immune response.

Several carcinogen-induced tumour lines, the 1509a fibrosarcoma, P-815-X2 mastocytoma and the SL2 lymphoma, as well as the spontaneous lymphoma L5178Y, when implanted subcutaneously in small numbers fail to produce death or a palpable tumour mass in a large proportion of syngeneic recipients (Table 1). To identify the number of injected cells capable of forming a tumour mass, medium containing 50 cells of L5178Y or P-815-X2 was placed in a soft agar gel which allows the formation of single cell clones¹⁰. Cloning efficiencies of up to 46% per inoculum were demonstrated after 3 week of culture (Table 1). The 1509a tumour cells were also shown, in repeated experiments, to grow well *in vitro*. Inocula of 10², 10³ and 10⁴ cells produced 5 × 10⁴, 10⁶ and 5 × 10⁶ tumour cells respectively, in 14 d. The capacity of the small tumour inocula to grow *in vivo* was demonstrated after intraperitoneal implantation of the 1509a, L5178Y and P-815-X2 (Table 1). In spite of the ability of 10² 1509a to grow intraperitoneally, intravenous

as well as subcutaneous challenge with that number of cells produced very few tumours (Table 1).

The rejection of the small subcutaneous inocula, however, could be prevented by the simultaneous intravenous injection of papain digest or 3 M KCl extracts of syngeneic tumour membrane (Table 2). Tumour frequency was 2.5–3.5 times higher than in untreated controls or controls receiving allogeneic membrane at the optimal dose of membrane extract.

The existence of memory was assessed in mice surviving the initial challenge of a small tumour inoculum by rechallenging with a larger tumour dose which was observed in preliminary experiments to result in 100% mortality. The latency and survival rate of mice which had previously rejected the L5178Y and P-815-X2 inocula was no different from that of normal control mice (Table 1). No difference was observed in the survival rates when mice were inoculated with tumour obtained from either *in vivo* peritoneal passage or *in vitro* tissue culture (Table 1). Similarly, mice which had rejected the 1509a in the primary challenge with the small inocula did not reduce either the growth rate or the mortality rate of the secondary challenge. It was also found that mice which had rejected 10² 1509a cells had the same tumour incidence as untreated controls when rechallenged with the same small inoculum (Table 1).

We examined the effects of suppression of the immune response on the incidence of tumours. A/J mice can be made resistant to 1509a by surgical removal of the primary tumour, and the subsequent simultaneous administration of suppressor T cells and 1509a to these immune mice results in enhanced tumour growth⁹. Suppressor T cells given with a small inoculum of 1509a in a primary challenge did not, however, result in any enhancement in the frequency of tumours or their growth rate. In addition, the simultaneous administration of suppressor

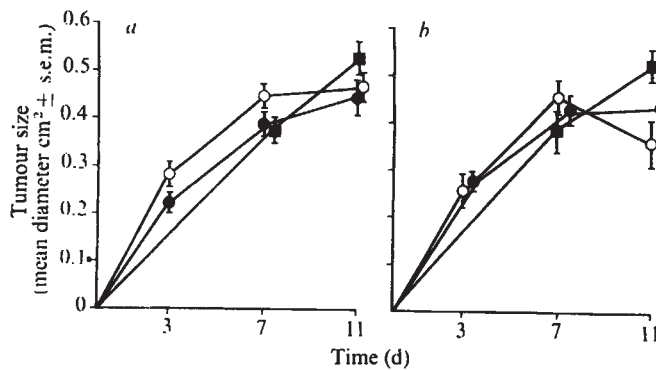


Fig. 1 Absence of memory after the rejection of a small 1509a tumour inoculum. *a*, Primary inoculation with 10^3 1509a subcutaneously was followed by a secondary challenge with 10^6 tumour cells 30 d later in the tumour-free mice. Suppressor T cells obtained from A/J thymus⁹ were given intravenously, simultaneously with the 10^6 inoculum in a second group of mice. Control mice did not receive a primary inoculation before 10^6 tumour challenge. The tumour size was measured with a Vernier caliper in two diameters at right angles. Observations were made on five mice in each experimental group. (●) 10^6 1509a secondary challenge after rejection in 10^3 primary challenge; (○) 10^6 secondary challenge after injection of a primary challenge with 10^3 1509a and 4×10^7 suppressor T cells; (■) 10^6 1509a primary challenge. *b*, Experimental protocol is identical to *a* except the primary inoculation is 10^3 1509a.

T cells and the small primary tumour inoculum did not enhance the growth rate of the secondary challenge of 1509a (Fig. 1).

We examined further the question of whether the rejection of small tumour inocula was a thymus-dependent adaptive immune phenomenon by subcutaneously implanting the 1509a and the L5178Y in their respective syngeneic hosts after adult thymectomy, lethal irradiation and bone marrow reconstitution. The frequency of tumours was slightly less in the immunodeficient mice than in the intact control mice, in spite of a much reduced survival rate when challenged with a large tumour dose (Table 3).

If the inoculation of a small number of tumour cells is an adequate model for an oncogenic mutation, the ability of immunodeficient or immunosuppressed mice to reject them suggests that surveillance is not a thymus-dependent adaptive immune phenomenon. This thesis is also supported by the additional evidence that the rejection is not associated with memory.

The ability of syngeneic tumour membrane extracts to

Table 2 Enhancement of tumour frequency after subcutaneous implantation of small tumour inocula by the intravenous injection of syngeneic tumour membrane preparations

Membrane preparation Tumour*	Amount† (µg)	Tumour inoculum	Tumour frequency‡ Number	%
L5178Y	0.1	5×10^4 L5178Y	7/16	44
L5178Y	1.0	5×10^4 L5178Y	12/23	52
L5178Y	10.0	5×10^4 L5178Y	6/16	37
—	0	5×10^4 L5178Y	7/35	20
1509a	1.0	5×10^4 L5178Y	2/12	17
1509a	0.1	10^3 1509a	4/10	40
1509a	1.0	10^3 1509a	7/10	70
1509a	10.0	10^3 1509a	5/10	50
—	0	10^3 1509a	2/10	20
L5178Y	1.0	10^3 1509a	0/5	0

*Papain digest of L5178Y membrane was prepared by the method of Maramatsu *et al.*²¹. The membrane digest was centrifuged at 25,000 r.p.m. for 10 min and placed on a Sephadex G-200 column. A single protein peak of low molecular weight was observed. The 3 M KCl membrane extract of 1509a was prepared by the method of Forbes *et al.*²².

†1.4 = 0.7142 mg ml⁻¹ at E₂₈₀

‡Tumour frequency was assessed 25 d after the implantation of the subcutaneous tumour inoculum. The data represent a total of three experiments for the L5178Y and two experiments for 1509a.

Table 3 Frequency of tumours and mortality of adult thymectomised, lethally irradiated and bone marrow reconstituted mice (AT × BM) challenged with small and large tumour inocula

Experiment	Tumour	Inoculum size	Recipient†	Tumour‡ frequency No.	%	Survival (d ± s.e.m.)
1.	L5178Y*	5×10^4 5×10^4	Normal DBA/2 AT × BM DBA/2	3/15 0/15	20 0	> 40 —
2.	L5178Y	10^6 10^6	Normal DBA/2 AT × BM DBA/2	19/19 14/14	100 100	35 ± 2§ 18 ± 1
3.	1509a	10^3 10^3 10^3 10^3 10^6	Normal A/J AT × BM A/J Normal A/J AT × BM A/J Normal A/J	0/5 0/5 2/5 1/5 5/5	0 0 40 20 100	— — — — —

*Cloning efficiency of the L5178Y was $47 \pm 5\%$ per inoculum.

†AT × BM mice received 800 rad irradiation 4 weeks after adult thymectomy followed by 2.5×10^7 washed bone marrow cells intravenously. Mice were challenged with tumour 3–4 weeks later.

‡Tumour frequency was assessed at 30 d after inoculation of tumour.

§*P* < 0.001

enhance tumour frequency suggests that the failure of tumour growth is due to a rejection mechanism with at least some specificity. It is also clear from our observations that one of the features characteristic of this mechanism is the rapidity of its response, acting well before the tumour has reached a critical size. Since tumour growth is very rapid and the time required to reach this critical size is theoretically very small, it suggests that the recognition mechanism is functionally optimal in the normal animal and capable of acting before an adaptive immune response can be generated. Such pre-existing specific immune mechanisms have been identified in two forms in mammals: naturally occurring antibodies^{11,12} and naturally occurring cytotoxic cells^{13–15}.

Natural antibodies with specificity for spontaneous or carcinogen-induced tumour, such as were used in the present study, have been found in normal mouse sera as well as sera from congenitally athymic (*nu/nu*) mice and are predominantly IgM^{12,16}. The role of IgM in anti-tumour immunity is not entirely clear, but it has been implicated in cell-mediated immunity to murine sarcoma virus¹⁷ and tumours induced by mammary tumour virus¹⁸.

Naturally occurring cytotoxic cells have been demonstrated with well defined specificities in *in vitro* assays of tumours immunity^{13–15} and are under genetic control of a strong H-2-linked factor in mice¹⁹. These cells have been partially characterised and bear neither B- nor T-cell membrane antigens^{13,14,20}. Keissling *et al.*²⁰ have also recently demonstrated a correlation between *in vitro* cytotoxicity by naturally occurring cytotoxic cells and their ability to produce *in vivo* neutralisation of an MSV lymphoma in a Winn assay. It is not clear at this time whether mice possess these effector cells in sufficient heterogeneity to account for the recognition of the diverse populations of tumours expected as a consequence of carcinogen-induced mutations, viral transformation, or tumour induction by other oncogenes. But the possibility that these effector cells are responsible for the rejection phenomenon we have observed must certainly be entertained.

We believe, therefore, that if immune surveillance of tumours exists, the possibility that it functions as a thymus-independent, non-adaptive phenomenon must be considered. Although the nature of this type of a surveillance mechanism is not immediately apparent, its identification may be of considerable importance to tumour biology and immunotherapy.

We thank Dr A. Sehon for his support and editorial comments, Mrs E. Shewchuck and Miss Motoko Nishikawa for technical assistance, and Mrs J. MacDonald for typing this manuscript.

This work was supported by the MRC of Canada, the

National Cancer Institute of Canada, the Children's Hospital Research Foundation and the NIH.

A. H. GREENBERG

Department of Pediatrics and Immunology,
Manitoba Institute of Cell Biology,
700 Bannatyne Avenue,
Winnipeg, Canada R3E 0V9

M. GREENE

Department of Immunology,
University of Manitoba,
Winnipeg, Canada

Received June 1; accepted September 29, 1976.

- ¹ Thomas, L., *Cellular and Humoral Aspects of the Hypersensitivity States* (edit. by Lawrence, H. S.), 529 (Cassell, London 1959).
- ² Burnet, M., *Immunological Surveillance* (Pergamon, London 1970).
- ³ Prehn, P. T., *Proc. 10th Can. Cancer Res. Conf. Honey Harbour, Ontario* (edit. by Scholefield, P. G.), 137 (University of Toronto Press, 1973).
- ⁴ Prehn, P. T., *Immune Surveillance: Proc. Brook Lodge Conf.* (edit. by Smith, P. T., and Landy, M.), (Academic, New York, 1975).
- ⁵ Rygaard, J., and Povlsen, C. O., *Transplant. Rev.*, **28**, 43–60 (1976).
- ⁶ Gillette, R. W., and Fox, A., *Cell. Immun.*, **19**, 328–335 (1975).
- ⁷ Stutman, O., *Science*, **183**, 534–536 (1974).
- ⁸ Skov, C. B., Holland, J. M., and Perkins, E. H., *J. natn. Cancer Inst.*, **56**, 193–195 (1976).
- ⁹ Fujimoto, S., Greene M. I., and Sehon, A., *J. Immun.*, **116**, 791–799 (1976).
- ¹⁰ Chu, M. Y., and Fischer, G. A., *Biochem. Pharmac.*, **17**, 753–760 (1968).
- ¹¹ Herberman, R. B., and Aoki, T., *J. exp. Med.*, **136**, 94–111 (1972).
- ¹² Martin, S. T., and Martin, W. J., *Int. J. Cancer*, **15**, 650–664 (1975).
- ¹³ Gomard, E., Leclerc, J. C., and Levy, J. P., *Nature*, **250**, 671–673 (1974).
- ¹⁴ Greenberg, A. H., and Playfair, J. H. L., *Clin. exp. Immun.*, **16**, 99–110 (1974).
- ¹⁵ Herberman, R. B., Nunn, M. B., and Lourin, D. H., *Int. J. Cancer*, **16**, 216–229 (1975).
- ¹⁶ Martin, W. J., and Martin, S. T., *Nature*, **249**, 564–565 (1974).
- ¹⁷ Lamon, E. W., Whitton, H. D., Skurzak, H. N., Andersson, B., and Liden, B., *J. Immun.*, **115**, 1288–1294 (1975).
- ¹⁸ Blair, P. B., Lane, M. A., and Mar, P., *J. Immun.*, **113**, 606–609 (1976).
- ¹⁹ Kiehl, R., Petryanyi, G., Klein, F., and Wigzell, H., *Int. J. Cancer*, **15**, 933–940 (1976).
- ²⁰ Kiehl, R., Petryanyi, G., Klein, F., and Wigzell, H., *Int. J. Cancer*, **17**, 275–281 (1976).
- ²¹ Maramatsu, T., Nathenson, S. G., Boyse, E. A., and Old, L. J., *J. exp. Med.*, **137**, 1256–1261 (1973).
- ²² Forbes, J. T., Yoshinobu, N., and Smith, R. T., *J. exp. Med.*, **141**, 1181–1186 (1975).

X chromosome activity in oocytes of kangaroo pouch young

In female eutherian mammals, inactivation of one of the two X chromosomes occurs randomly from cell to cell in somatic tissues¹ but in oocytes (meiotic germ cells) X inactivation does not occur for at least three loci. Oocytes of adult XX mice possess twice as much activity as those of XO mice for the sex-linked enzyme, glucose-6-phosphate dehydrogenase (G6PD)², hypoxanthine guanine phosphoribosyl transferase (HGPRT)³, and phosphoglycerate kinase (PGK)⁴. Furthermore, oocytes of adult⁵ and foetal⁶ human females, heterozygous for G6PD allelic genes, show activity for both fast and slow electrophoretic forms of the enzyme. In female kangaroos (Marsupialia: Macropodinae) inactivation of the paternally derived X chromosome occurs in cells of various somatic tissues⁷ but the question of X chromosome activity in oocytes has not been resolved. We have now examined G6PD expression in ovaries of pouch young known to be heterozygous for the fast and slow forms of the enzyme, and show that only a single X chromosome—the maternally derived one—is active in ovarian cells.

Inheritance of G6PD is sex linked in kangaroos⁸. Crosses between wallaroos (*Macropus robustus robustus*: G6PD-F) and euros (*M. r. erubescens*: G6PD-S) produce fertile female hybrids heterozygous for G6PD type. Electrophoresis was carried out on ovaries from six heterozygous pouch young aged between 53 and 60 d *post partum*. This age class has the greatest number of oocytes during ovarian development of *Macropus eugenii*⁹ and we assume the same is the case for *M. robustus*. Graphical reconstruction of cell and tissue distribution from serial sections of an ovary from a 53-d-old pouch young showed that primary oocytes accounted for about 70% of the total ovarian volume, a far higher proportion than at any stage of human ovarian development¹⁰.

Most of the oocytes were in zygotene or pachytene of prophase of meiosis I. Specific staining¹¹ of frozen sections of an ovary from another heterozygous pouch young aged 57 d showed that oocytes had a high level of G6PD activity. Although we have no quantitative data on this point, it seems unlikely that activity levels differ greatly from those in human oocytes at comparable stages of differentiation.

If both alleles are functional in oocytes heterozygous for G6PD and the enzyme molecule is a dimer, then electrophoretic bands should be present in the fast and slow positions together with a stronger intermediate or heteropolymer band. In the ovarian extracts examined, however, only a single band was found occupying the maternal position regardless of the direction of the cross, and there was never any trace of a band in the paternal position (Fig. 1). A brief reference to two of these heterozygotes was made earlier¹².

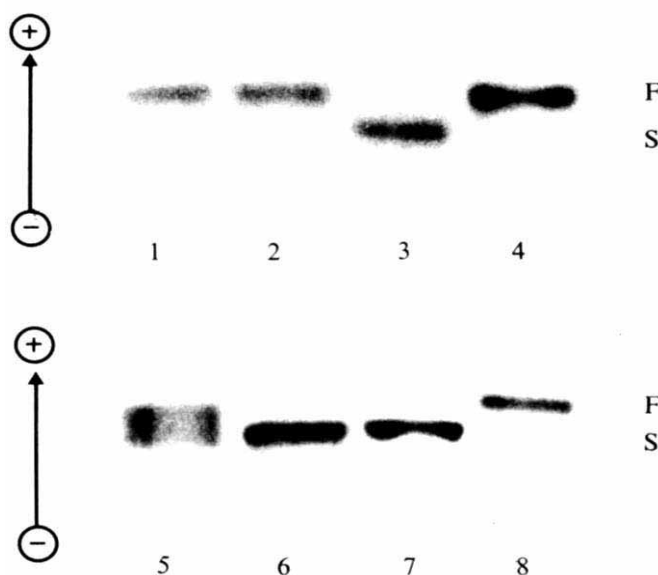


Fig. 1 G6PD electrophoretic phenotypes (F, fast; S, slow) from ovarian and other tissue homogenates of heterozygous pouch young. In the F/S heterozygote (age: 55 d) and the S/F heterozygote (56 d) the allele of maternal origin is referred to first. The S/S homozygote is used as a control. 1, F/S brain; 2, F/S ovary; 3, S/S ovary; 4, F/S uterus; 5, 1:1 mixture of F/S and S/F kidneys; 6, S/F ovary; 7, S/F oesophagus; 8, F/S oesophagus. Channels 1–4 and 5–8 represent separate gels. Samples were homogenised in an equal volume of lysing solution¹³ and applied on to Cellologel (Chemotron). Electrophoresis was carried out in 0.1 M lithium borate–0.0024 M EDTA buffer, pH=9.0, for 1.5 h with a voltage gradient of 14 V cm⁻¹. The gel stain was that of Rattazzi *et al.*¹¹.

In view of the preponderance of oocytes in the ovaries examined and their relatively high levels of G6PD activity, it seems reasonable to conclude that X inactivation of the allele of paternal origin occurs in these cells. At our level of detection it is most unlikely that gene product, forming bands in paternal and intermediate positions, could go undetected. We consider, therefore, that inactivation of the paternal G6PD allele is a feature not only of female kangaroo somatic cells but also of their early primary oocytes.

In human foetal ovaries, results of electrophoresis of G6PD heterozygotes were interpreted as evidence for a reactivation of the inactive X chromosome of germ line cells near the time of entry to meiosis¹³, rather than at the differentiation of primordial germ cells into oogonia as had previously been suggested¹⁴. Foetal ovaries with numerous

zygotene and pachytene oocytes gave a distinct heteropolymer band but in younger specimens with few leptotene oocytes, the heteropolymer band was weak or not detectable at all¹³. Ovaries of heterozygous kangaroo pouch young, in which zygotene and pachytene oocytes were predominant, gave no trace of a heteropolymer band, indicating that if reactivation of the paternal allele occurs in kangaroo oocytes it occurs at a later stage of oogenesis. We are at present investigating adult heterozygote ovaries to establish whether reactivation has occurred by the dictyate oocyte stage.

That reactivation of the paternal allele for G6PD does occur at some time during oogenesis (or during embryogenesis) is supported by kangaroo pedigree data¹⁵. Progeny derived from G6PD heterozygotes may inherit either allele in an active form, suggesting that both X chromosomes are potentially active, at least at this locus. We assume, as have other workers, that activity states at the G6PD locus represent activity states at most if not all X-linked loci so that inactivation of the paternal allele for G6PD is indicative of inactivation of the paternal X chromosome. The patterns of transcription at other X-linked loci in kangaroo oocytes need to be elucidated to validate the supposition.

Kangaroo pachytene oocytes possess an XX bivalent which is isopycnotic with the autosomal bivalents. Air-dried preparations showed no differential condensation of either homologue or regions of homologues. There is thus no cytological evidence of a difference in genetic activity between the maternal and paternal X chromosomes of the kangaroo sex bivalent.

The mechanisms which control the maintenance of X inactivation early in oocyte differentiation and the onset of reactivation at some later stage are not known and are worthy of investigation in the more general context of regulation of activity states in chromosomes.

We thank Drs D. W. Cooper and D. A. Briscoe for discussions and for reading the manuscript, Mr Ron Moore for technical assistance and we also acknowledge financial support from the Australian Research Grants Committee and Macquarie University Research Fund.

P. G. JOHNSTON
E. S. ROBINSON
G. B. SHARMAN

School of Biological Sciences,
Macquarie University,
North Ryde, Australia 2113

Received September 8; accepted September 28, 1976.

- ¹ Lyon, M. F., *Nature*, **190**, 372-373 (1961).
- ² Epstein, C. J., *Science*, **163**, 1078-1079 (1969).
- ³ Epstein, C. J., *Science*, **175**, 1467-1468 (1972).
- ⁴ Kozak, L. P., McLean, G. K., and Eicher, E. M., *Biochem. Genet.*, **11**, 41-47 (1974).
- ⁵ Gartler, S. M., Liskay, R. M., Campbell, B. K., Sparkes, R., and Gant, N., *Cell Differentiation*, **1**, 215-218 (1972).
- ⁶ Gartler, S. M., Liskay, R. M., and Gant, N., *Expl Cell Res.*, **82**, 464-466 (1973).
- ⁷ Cooper, D. W., et al., in *Isozymes, III Developmental Biology* (edit. by Markert, L.), 559-573, (Academic, New York, 1975).
- ⁸ Richardson, B. J., Czuppon, A., and Sharman, G. B., *Nature new biol.*, **230**, 154-155 (1971).
- ⁹ Alcorn, G. T., thesis, Macquarie Univ. (1975).
- ¹⁰ Baker, T. G., *Proc. R. Soc. B158*, 417-433 (1963).
- ¹¹ Rattazzi, M. C., Bernini, L. F., Fiorelli, G., and Manucci, P. M., *Nature*, **213**, 79 (1967).
- ¹² Johnston, P. G., and Sharman, G. B., *Aust. J. Biol. Sci.*, **28**, 567-574 (1975).
- ¹³ Gartler, S. M., Andina, R., and Gant, N., *Expl Cell Res.*, **91**, 454-457 (1975).
- ¹⁴ Ohno, S., *2nd Int. Conf. Congenital Malformations*, 36-40 (New York, 1963).
- ¹⁵ Sharman, G. B., in *Cytotaxonomy and Vertebrate Evolution* (edit. by Chiarelli, A. B., and Capanna, E.), 485-530 (Academic, London, 1973).

pattern of cell growth, a limited life span *in vitro* and are not tumorigenic *in vivo*, are then converted into cells with a hereditary random pattern of cell growth, an ability to grow continuously in culture and to form tumours *in vivo*. Studies with a variety of polycyclic hydrocarbons and other chemicals have shown a correlation between the frequency of colonies with a hereditary random pattern of cell growth (transformed colonies) and the degree of carcinogenicity in animals^{2,6,7}.

The analysis of cell transformation by benzo[a]pyrene, a prototype and most common carcinogenic polycyclic hydrocarbon, has indicated that transformation by this carcinogen is a one-hit event³. This one-hit response and the observation that carcinogens can bind to cellular DNA^{4,9} suggest that transformation may result from a single alteration in the genetic constitution of the treated cell. Carcinogens may, therefore, transform cells as a result of a somatic mutation. It has indeed been found that chemical carcinogens can produce somatic mutations at various genetic loci¹⁰⁻¹³ and that there is a relationship between the degree of carcinogenicity and mutagenicity with different chemicals¹²⁻¹³. The relationship between carcinogenesis and mutagenesis in these studies was not, however, tested with the same cells, nor was mutagenesis by environmental carcinogens tested with normal diploid cells.

To elucidate the mechanism of carcinogenesis, it is important to determine the number of the same or different genes which may be involved in cell transformation. The frequency of cell transformation and mutagenesis of specific loci can both be tested *in vitro*. We have, therefore, now carried out experiments with normal diploid cells to determine whether the ratio between transformation and mutation of a specific locus is small, about 10, or large, more than 10³. We have determined mutation and transformation frequencies in primary cultures of normal golden hamster embryo cells treated with the carcinogenic hydrocarbon benzo[a]pyrene and its metabolite 7,8-trans-dihydrodiol¹⁴. The frequency of transformed colonies has been compared with the frequency of mutation for ouabain resistance, a mutation associated with the surface membrane Na⁺/K⁺ ATPase^{16,17}. This mutation was chosen because of its low spontaneous frequency and the ease with which it can be detected in normal diploid cells¹⁷. The cells were treated in the same way to test for the frequencies of transformation and ouabain-resistant mutants.

The results have shown that benzo[a]pyrene and its 7,8-dihydrodiol can induce both transformation and mutagenesis for ouabain resistance in normal diploid cells. The ratio between transformation and mutagenesis for ouabain resistance was about 20:1 with both hydrocarbons (Table 1). Other studies with normal hamster embryo cells treated with benzo[a]pyrene, have indicated a transformation-to-mutation ratio of about the same order of magnitude (J. C. Barrett and P. O. P. Ts'o, personal communication). Since ouabain resistance is presumably due to mutation at one locus¹⁷, it can be suggested that transformation as measured by the appearance of colonies with a random pattern of cell growth has a target equivalent to about 20 such genes. There may thus be this number of the same or different genes, the mutation of any one of which is the single hit required for transformation. Further studies with other genes can determine whether some genes show a lower or higher transformation-to-mutation ratio and whether this ratio changes with different types of carcinogens. The number of genes involved in transformation may be even smaller than about 20, if the genes for transformation are located at hot spots¹⁸ which have a higher mutation frequency. It is of interest that genes which determine the production of altered heavy chains of immunoglobulin in mouse myeloma cells also have a high mutation frequency^{19,20}, of the same order of magnitude as the frequency of cell transformation.

Mutagenesis and transformation of normal cells by chemical carcinogens

TREATMENT of normal cells¹⁻⁸ or some established cell lines⁹ in culture with chemical carcinogens can result in cell transformation and malignancy. Cells which have an oriented

Table 1 Transformation and mutagenesis by benzo[a]pyrene and its 7,8-dihydrodiol

Hydrocarbon	Concentration ($\mu\text{g ml}^{-1}$)	Cloning efficiency (%)	No. of transformed colonies per no. of colonies counted	No. of transformed colonies per 10^5 colony-forming cells	No. of ouabain resistance colonies	A/B
				(A)	(B)	
None	—	12.5	0/6,736	—	0.04	—
Benzo[a]pyrene	1	1.2	16/3,753	4,300	213	20
7,8-dihydrodiol	0.1	2.6	25/5,309	4,700	193	24
	0.3	1.3	18/2,203	8,200	451	18
	1.0	0.9	20/3,183	6,300	419	15

Three- to five-day-old primary normal golden hamster embryo fibroblasts were seeded at 1.5×10^5 cells in 8 ml Eagle's medium with a four-fold concentration of amino acids and vitamins (H-21, Grand Island) with 10% foetal calf serum in 100 mm tissue culture (Falcon) Petri dishes. One day after seeding, the cells were treated for 3 d with benzo[a]pyrene or its 7,8-dihydrodiol metabolite dissolved in a final concentration of 0.5% acetone. The cells were then washed with medium, dissociated with trypsin solution and seeded to determine the frequency of transformed and ouabain-resistant colonies. To test for transformation the treated cells were seeded as described¹⁻³ and the number of colonies counted 10 d after seeding. To test for ouabain resistance¹⁴⁻¹⁷, the carcinogen-treated cells were seeded at 10^5 cells per 50 mm Petri dish. Two days later ouabain was added at a final concentration of 1 mM and the number of colonies counted after 18 d. In all experiments colonies were counted after staining with Giemsa. The data presented are the total results of four separate experiments per point. At the plateau concentration, there was up to about a fourfold variation in the assays for transformed and ouabain-resistant colonies. The heritability of the ouabain resistance was shown with five isolated ouabain-resistant colonies, all of which remained resistant to the drug after 10–12 d growth in the absence of ouabain.

We suggest that cell transformation is due to a mutation and that this mutation can occur in one out of a small number of the same or different genes. Cell transformation by chemicals resulting in a hereditary random pattern of cell growth is an initial step in the process that can result in the formation of tumours^{4,5}. In addition to the mutation resulting in transformation, it also seems that there can then be other genetic changes such as the selection of specific genes that control the ability of cells to form a tumour²¹⁻²³.

This study was supported by a contract with the NCI, US PHS.

ELIEZER HUBERMAN
RIVKA MAGER
LEO SACHS

Department of Genetics,
Weizmann Institute of Science,
Rehovot, Israel

Received June 14; accepted October 4, 1976.

- Berwald, Y., and Sachs, L., *Nature*, **200**, 1182–1184 (1963).
- Berwald, Y., and Sachs, L., *J. natn. Cancer Inst.*, **35**, 641–661 (1965).
- Huberman, E., and Sachs, L., *Proc. natn. Acad. Sci. U.S.A.*, **56**, 1123–1129 (1966).
- Sachs, L., *Curr. Top. in dev. Biol.*, **2**, 129–150 (Academic, New York, 1967).
- Huberman, E., Salzberg, S., and Sachs, L., *Proc. natn. Acad. Sci. U.S.A.*, **59**, 77–82 (1968).
- DiPaolo, J. A., Nelson, R. L., and Donovan, P. J., *Nature*, **235**, 278–281 (1972).
- Heidelberger, C., *Adv. Cancer Res.*, **18**, 317–366 (1973).
- Brookes, P., and Lawley, P. D., *Nature*, **202**, 781–784 (1964).
- Kuroki, T., and Heidelberger, C., *Cancer Res.*, **31**, 2168–2176 (1971).
- Kao, F., and Puck, T. T., *Proc. natn. Acad. Sci. U.S.A.*, **60**, 1275–1281 (1968).
- Kao, F., and Puck, T. T., *J. Cell Physiol.*, **78**, 139–144 (1971).
- Ames, B. N., Durston, W. E., Yamasaki, E., and Lee, F. D., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 2281–2285 (1973).
- McCann, J., and Ames, B. N., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 950–954 (1976).
- Huberman, E., and Sachs, L., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 188–192 (1976).
- Huberman, E., Sachs, L., Yang, S. K., and Gelboin, H. V., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 607–611 (1976).
- Baker, R. M., et al., *Cell*, **1**, 9–21 (1974).
- Mankovitz, R., Buchwald, M., and Baker, R. M., *Cell*, **3**, 221–226 (1974).
- Benzer, S., *Proc. natn. Acad. Sci. U.S.A.*, **47**, 403–415 (1961).
- Baumal, R., Birshtein, B. K., Coffino, P., and Scharff, M. D., *Science*, **182**, 164–166 (1973).
- Preud'homme, J.-L., Birshtein, B. K., and Scharff, M. D., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1427–1430 (1975).
- Hitotsumachi, S., Rabinowitz, Z., and Sachs, L., *Nature*, **231**, 511–514 (1971).
- Sachs, L., *Harvey Lectures*, **68**, 1–35 (Academic, New York, 1974).
- Bloch-Shtacher, N., and Sachs, L., *J. Cell Physiol.*, **87**, 89–100 (1976).

Dipole potential measurements in asymmetric membranes

We report here detection of two kinds of asymmetry in planar lipid bilayers which may bear on the problem of asymmetry in biological membranes. One kind of asymmetry arises as a consequence of an asymmetric distribution in the bilayer of lipids with charged polar head groups, and the other from an asymmetric distribution of lipids with different but neutral polar

head groups. Several studies indicate that the lipids of the red cell membrane are asymmetrically distributed. In the erythrocyte membrane the outer half of the bilayer consists predominantly of neutral lipids, whereas the inner half contains all the negatively charged phospholipid phosphatidylserine¹⁻³. In nerve axons such lipid asymmetry has not been demonstrated directly but can be inferred from surface charge measurements. These measurements indicate that at least in the neighbourhood of the sodium channel, the negative charge density is greater in the outer monolayer of the membrane^{4,5}.

Recently it has been shown that lipid redistribution between the two halves of the bilayer is a very slow process⁶⁻⁸, thus indicating that membrane asymmetry can be a stable characteristic of biological membranes. A natural consequence of lipid asymmetry in membranes is a difference in surface potentials. In appropriate conditions this difference can have a major role in controlling ion transport^{9,10}. A surface potential may arise either from ionised groups or from oriented dipoles at the membrane surface or both, and there were indications that the dipole contribution is such that the membrane interior is > 300 mV positive with respect to the external solutions¹⁰. Surface charge usually makes the membrane more negative with respect to the solutions. In previous work we showed that surface charge asymmetry in bilayer membranes can be detected and measured using the nonactin- K^+ complex as a probe⁷. Here we report a similar measurement of the dipole potential difference between two lipids with different but neutral polar head groups. The potential measured in this case does not arise from surface charge, but from a difference in dipole moment per unit area in the two lipid monolayers which form the membrane.

The technique used to form asymmetric bilayers by apposition of the two separate monolayers has been previously described^{7,11}. In the present experiments one monolayer was formed from bacterial phosphatidylethanolamine (PE, Supelco, Inc.) and the other from 1,3 Diolein (GDO, Applied Science Lab., Inc., 99 per cent+) or bovine phosphatidylserine (PS, Supelco). The aqueous solutions in both compartments were equal and consisted of unbuffered ($\text{pH} \sim 6$) KCl solutions of indicated concentration. Nonactin (E.R. Squibb) was added to both compartments in concentrated ethanolic solution; the dilution was usually by a factor of 10^3 . Electrical measurements were made using Ag/AgCl electrodes in a four electrode system¹². Current-voltage curves were directly recorded on an X-Y recorder. The rate of change of voltage was 10 mV s^{-1} or less to avoid hysteresis. All the experiments were made at $20 \pm 2^\circ \text{C}$.

Figure 1 shows a steady-state current-voltage curve of an asymmetric bilayer membrane in which one monolayer is made of GDO and the other of PE. The voltage is defined as positive

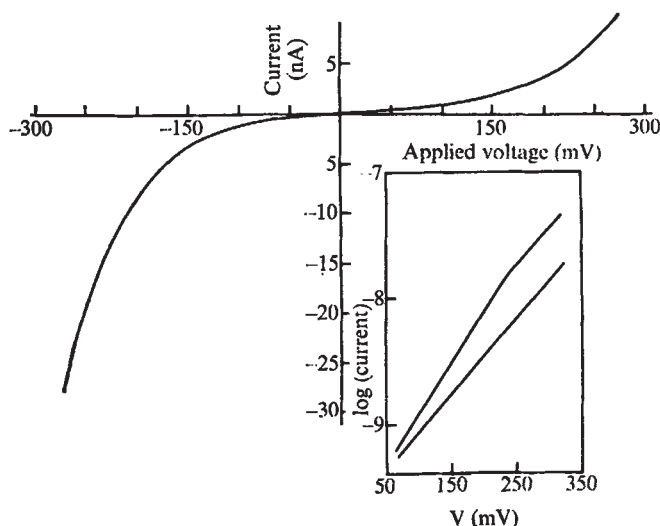


Fig. 1 Steady-state current-voltage curve for an asymmetric membrane made of GDO and PE. Nonactin concentration 2×10^{-8} M. 1 M KCl unbuffered ($pH \approx 6$). Temperature $21^\circ C$. Positive current is defined as flowing from the high dipole potential side (PE) to the low dipole potential side (GDO). The insert shows the data plotted on a logarithmic scale. Note the change in slope for current flowing from the GDO side to the PE side for negative potential in the upper curve. The lower curve is for positive potentials and current flowing from the PE to the GDO side.

when current flows from the PE side across the membrane to the GDO side. The rectification shown by the current-voltage curve is consistent with a more positive dipole potential on the PE side. This rectification arises as a consequence of the actual energy barrier to transport inside the membrane, but it is easier to understand with the trapezoidal approximation shown in Fig. 2. We have previously shown that for large applied voltages a trapezoidal approximation of the actual barrier profile is adequate to describe the steady state current-voltage curves in bacterial PE membranes¹³ as well as those found in asymmetric membranes made of bacterial PE and phosphatidylserine⁷. Figure 2a shows the trapezoidal approximation for the energy barrier profile expected for an asymmetric membrane with a difference in dipole potential, $\Delta\phi$, across it. Studies made in

both monolayers¹⁰ and symmetric bilayers^{10,14} indicate that $\phi\Delta$ between GDO and PE is ~ 100 mV, the interior of a PE membrane being more positive. As shown in Fig. 2b the rate limiting energy for the currents in the positive direction, ϕ_{PE} , is reduced by a positive voltage, nV , where n is the fractional distance to the corner. On the other hand, the rate-limiting energy for the current in the negative direction, $\phi_{GDO} + \Delta\phi$ in Fig. 2c, is reduced by a negative voltage, $(1-n)V$. Accordingly, the currents in the positive and negative potentials at large voltages must be proportional to $\exp(neV/kT)$ and to $\exp[e(1-n)V/kT]$ respectively, and we would expect a difference in the functional dependence of the current on voltage for different sign of applied voltage like the one shown in Fig. 1. Differences in dipole potential can be calculated with ~ 5 -mV accuracy from the differences in functional dependence of the current on voltages of different polarities. The insert of Fig. 1 shows the logarithm of the current against voltage. The upper curve in the insert represents current flowing from the GDO side to the PE side and the lower curve the reverse current. Both curves diverge until a voltage at which the upper curve changes its slope and becomes parallel to the lower curve. At this point in voltage the external applied potential has compensated exactly for the differences in dipole potential between the two lipids. The calculated difference in dipole potential is 107 mV in 0.02 M and 101 mV in 1 M KCl. It is important to point out that no rectification in the current-voltage curve is expected from a triangular barrier, a shape frequently assumed in the study of carrier mediated transport.

We previously reported⁷ asymmetry in membranes formed with one monolayer of PE and the other of PS. Figure 3 shows the change with ionic strength in the surface potential difference of both PE-PS membranes and PE-GDO membranes. The PE-PS membranes show an asymmetry which depends on ionic strength as predicted on the Gouy-Chapman theory. The asymmetry of PE-GDO membranes is, however, independent of ionic strength. This observation is in agreement with the fact that PE and GDO are neutral at the pH used ($pH \sim 6$). This lack of change in surface potential with a change in ionic strength is an important characteristic of surface potentials arising from dipoles rather than surface charge.

Differences in dipole potential between different lipids have been inferred by measuring and comparing the specific conductance of symmetrical membranes in the presence of nonactin¹⁵ or lipophilic ions⁹. In the absence of space charge and

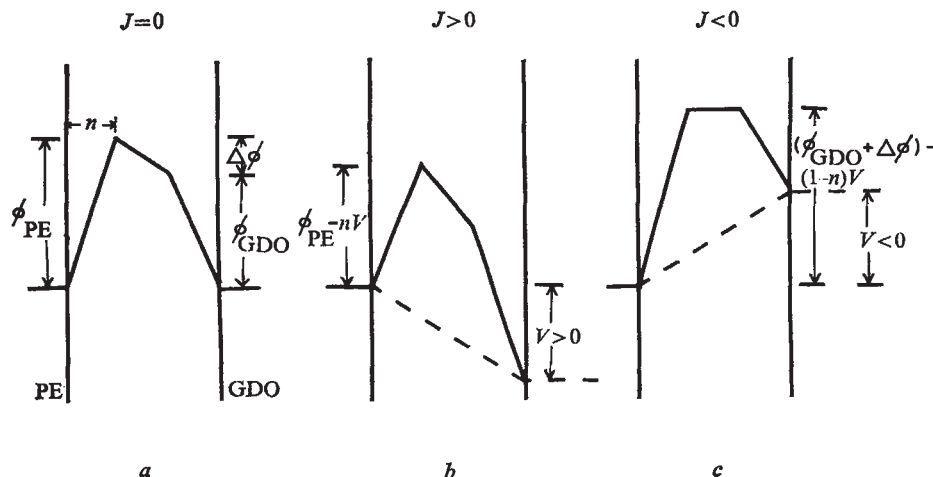


Fig. 2 Schematic representation of the energy barrier of an asymmetric membrane made of GDO-PE. ϕ_{PE} is the barrier height 'seen' by the nonactin-K complexes at the PE side. $\phi_{GDO} + \Delta\phi$ is the height of the barrier 'seen' by the nonactin-K⁺ complexes at the GDO side. $\Delta\phi$ is the difference in dipole potential between both lipids. n , the fractional distance to the corner, is 0.32 ± 0.03 . a, The expected barrier profile obtained by superposition of the dipole potential profiles for a GDO and PE monolayer. $\phi_{PE} = \phi_{GDO} + \Delta\phi_0$ and therefore there is no net flow of current at 0 potential. b, When a positive voltage V is applied, ϕ_{PE} is reduced by nV and the current flux J is proportional to $\exp(neV/kT)$ for large potentials ($V > 75$ mV). c, The negative potential applied exactly compensates for the differences in dipole potential. By using similar triangles it is easily shown that $\Delta\phi = V(1-2n)$. Here the barrier height ($\phi_{GDO} + \Delta\phi$) is reduced by $(1-n)V$ and therefore J is proportional at large voltages to $\exp[e(1-n)V/kT]$. If $V > \Delta\phi/(1-2n)$ then the height of the right hand corner of the barrier will become the current limiting step.

kinetic limitations from the formation of complexes at the membrane surface, the present method provides a way to measure such a difference directly, and does not depend on the nonactin concentration because it measures a difference in shape and not in magnitude. It also allows us to monitor

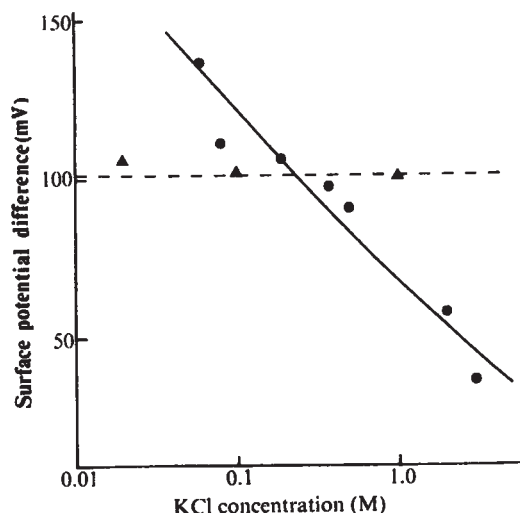


Fig. 3 Surface potential difference as a function of ionic strength. ●, Asymmetry potentials of membranes formed from one monolayer of phosphatidylserine and one of phosphatidylethanolamine. The dependence of surface potential on ionic strength is very nearly as predicted from the Gouy-Chapman equation with a surface charge density of 1 electronic charge per $75 \text{ \AA}^2$ (shown by the solid line). ▲, Asymmetry potentials of membranes formed from one monolayer of phosphatidylethanolamine and one of glycerol dioleate. The lack of dependence of surface potential on ionic strength is in agreement with a potential arising from dipoles rather than surface charge. Temperature $21 \pm 1^\circ \text{C}$.

changes in membrane asymmetry with time for membranes formed with two neutral lipids as long as they have a difference in dipole potential. Detection of 'flip-flop' in lipids in planar asymmetrical bilayer membranes has been restricted to membranes containing charged lipids^{16,7}. So far we have been unable to detect any lipid redistribution in membranes made of GDO and PE; this confirms other results⁶⁻⁸. It is of interest to note that we have detected no asymmetry in the current-voltage curves of membranes made of glycerolmonooleate and PE, even though the differences in dipole potential are $\sim 120 \text{ mV}$ (ref. 15). A possible explanation is that glycerolmonooleate distributes itself very rapidly between the two monolayers.

We have interpreted the asymmetry of our voltage-current curves as arising solely from differences in surface potential between two lipids. We believe this to be the case for two reasons. First, the asymmetry arising from charged lipids can be unambiguously attributed to a change in electric field across the membranes because of its dependence on ionic strength. Second, our neutral lipid results are consistent with values of surface potential derived from both monolayer and symmetric membrane studies^{10,13,15}.

As previously mentioned, lipid asymmetries are apparently present in many biological membranes. The differences in surface potential brought by the asymmetric lipid distribution can play an important role in affecting various charged structures in the membrane, for example the gating charges in nerves. Our work provides a means of measuring asymmetric potentials in membranes of well defined composition. Our results thus enable us to estimate how lipid composition may affect all electrical functions of biological membranes.

This work was supported in part by the NIH and The Louis Block Fund, The University of Chicago. We would like to thank Mr W. Koroshetz for help with some of the experiments.

RAMON LATORRE

Department of Pharmacological and Physiological Sciences,
The University of Chicago,
Chicago, Illinois 60637

JAMES E. HALL

Department of Physiology and Pharmacology,
Duke University,
Durham, North Carolina 27710

Received July 9; accepted September 27, 1976.

- ¹ Zwaal, R.F.A., Roelofsen, B., and Colley, C. M., *Biochim. biophys. Acta.*, **300**, 159-182 (1973).
- ² Gordeky, S. E., Marinetti, G. V., and Love, R., *J. Membr. Biol.*, **20**, 111-132 (1975).
- ³ Fisher, K. A., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 173-177 (1976).
- ⁴ Chandler, W. K., Hodgkin, A. L., and Meves, H., *J. Physiol., Lond.*, **180**, 821-836 (1965).
- ⁵ Hille, B., Woodhull, A. M., and Shapiro, B. I., *Proc. Trans. R. Soc. Lond.*, **B270**, 301-318 (1975).
- ⁶ Rothman, J. E., and Dawidowicz, E. A., *Biochemistry*, **14**, 2809-2816 (1975).
- ⁷ Hall, J. E., and Latorre, R., *Biophys. J.*, **15**, 99-103 (1976).
- ⁸ Parnansky, M., and Lange, Y., *Nature*, **259**, 420-421 (1976).
- ⁹ Szabo, G., *Nature*, **252**, 47-49 (1974).
- ¹⁰ Haydon, D. A., *Ann. N. Y. Acad. Sci.*, **264**, 2-16 (1975).
- ¹¹ Montal, M., *Biochim. biophys. Acta.*, **298**, 750-754 (1973).
- ¹² Eisenberg, M., Hall, J. E., and Mead, C. A., *J. Membr. Biol.*, **14**, 143-176 (1973).
- ¹³ Hall, J. E., Mead, C. A., and Szabo, G., *J. Membr. Biol.*, **11**, 75-97 (1973).
- ¹⁴ Szabo, G., and Eisenman, G., *Biophys. Soc. Abstr.*, **13**, 175a (1973).
- ¹⁵ Hlaky, S. B., and Haydon, D. A., *Biochim. biophys. Acta.*, **318**, 464-468 (1973).
- ¹⁶ Sherwood, D., and Montal, M., *Biophys. J.*, **15**, 417-439 (1975).

The effects of antipsychotics on the turnover rate of GABA and acetylcholine in rat brain nuclei

SEVERAL lines of indirect evidence indicate that the nigrostriatal and mesolimbic dopaminergic systems interact with acetylcholine (ACh)^{1,2,3} and γ -aminobutyric acid⁴ (GABA)-secreting neurones. The reciprocal relationships among these neurones have not, however, been established with precision. Nucleus caudatus and nucleus accumbens are innervated by dopaminergic axons which have their cell bodies located in substantia nigra and ventral tegmental area⁵, respectively. These nuclei contain high concentrations of ACh⁶ and GABA⁷ and a high activity of cholineacetyltransferase (CAT)⁶ and glutamic acid decarboxylase⁷ (GAD). The nucleus accumbens and nucleus caudatus are connected with globus pallidus and substantia nigra which also contain GABA⁷, ACh⁶ and their synthesizing enzymes. Antipsychotics, including chlorpromazine, clozapine and haloperidol increase the turnover rate of dopamine (DA) in the nucleus caudatus and nucleus accumbens⁸. Clozapine, a non-cataleptogenic antipsychotic, preferentially increases the DA turnover in the nucleus accumbens whereas chlorpromazine and haloperidol, two cataleptogenic antipsychotics show a preferential effect on the nucleus caudatus⁸.

We here show that the action of clozapine, chlorpromazine and haloperidol on the turnover rates of ACh and GABA in the nucleus caudatus, nucleus accumbens, globus pallidus and substantia nigra allows for a qualitative differentiation between the biochemical effects elicited by clozapine and cataleptogenic antipsychotics. The results not only shed some light on the possible functional interdependence among GABA, ACh and DA in neuronal systems of single brain nuclei but also they provide some biochemical information which may allow us to predict a pattern of specific neurotransmitter involvement in the genesis of the cataleptogenic effects of antipsychotics.

The turnover time of ACh in nucleus caudatus, substantia nigra, globus pallidus and nucleus accumbens was

measured by infusing phosphoryl (d_4) choline ($15 \mu\text{mol kg}^{-1} \text{ min}^{-1}$) at a constant rate, intravenously, for 9 min (ref. 9) and killing the rats at the end of the infusion with a high intensity microwave beam focused on the skull¹⁰. The brain was sliced into $400 \mu\text{m}$ thick slices and single brain nuclei were punched out as reported earlier⁸. Haloperidol ($10 \mu\text{mol kg}^{-1}$) clozapine ($30 \mu\text{mol kg}^{-1}$) or chlorpromazine ($20 \mu\text{mol kg}^{-1}$) were injected intraperitoneally 51 min before beginning the infusion with phosphoryl (d_4) choline. The deuterium enrichment of choline and ACh was measured by mass fragmentography in various brain nuclei^{8,9}. From these data the time needed to synthesise the ACh present in each nucleus (turnover time) was calculated¹¹. The estimation of GABA turnover time was performed by killing the rats at various times after the constant rate intravenous infusion of D-glucose uniformly labelled with ^{13}C ($50 \mu\text{mol kg}^{-1} \text{ min}^{-1}$ for 10 min). Rats were killed with a microwave beam at 2, 3.5 and 5 min after the end of infusion and various brain nuclei were dissected stereomicroscopically as described above. The GABA content was measured by mass fragmentography as reported earlier¹², the turnover time of GABA was calculated from the percentage incorporation of ^{13}C into glutamate and GABA¹³. Haloperidol ($4 \mu\text{mol kg}^{-1}$, intraperitoneally) and clozapine ($30 \mu\text{mol kg}^{-1}$, intraperitoneally) were injected 40 and 20 min respectively, before the infusion with ^{13}C -glucose.

Haloperidol, which is virtually devoid of anticholinergic effects¹⁴, reduced the ACh turnover time in the nucleus caudatus and nucleus accumbens, whereas clozapine, which is a potent muscarinic receptor blocker¹⁴⁻¹⁶ failed to change the turnover of ACh in these two nuclei (Table 1). The results obtained with haloperidol and chlorpromazine further support the current belief that nigrostriatal and mesolimbic dopaminergic synapses inhibit the metabolism of ACh in cholinergic postsynaptic-interneurons³. When dopaminergic receptors were blocked by these two cataleptogenic antipsychotics the metabolism of ACh in nucleus caudatus and nucleus accumbens increased (Table 1). Perhaps the decrease of ACh turnover time suggests that the activity of intrinsic cholinergic interneurons of nucleus caudatus and nucleus accumbens was increased. Our findings with ACh turnover time in the nucleus accumbens of rats are at variance with results obtained with push-pull cannulae perfusion of limbic and striatal areas of cats. These results show that cataleptogenic antipsychotics increase the efflux of ACh from striatum but not from limbic areas¹⁷. Since clozapine and haloperidol block the inhibition of striatal ACh metabolism caused by apomorphine¹⁸ it can be inferred that clozapine and haloperidol can block the postsynaptic, dopamine receptors in the nucleus caudatus¹⁶. Moreover, when clozapine was injected in rats receiving haloperidol the metabolism of striatal ACh failed to increase¹⁶. No direct explanation can be given for the different action of clozapine and cataleptogenic antipsychotics on the ACh turnover rate in the nucleus caudatus and nucleus accumbens.

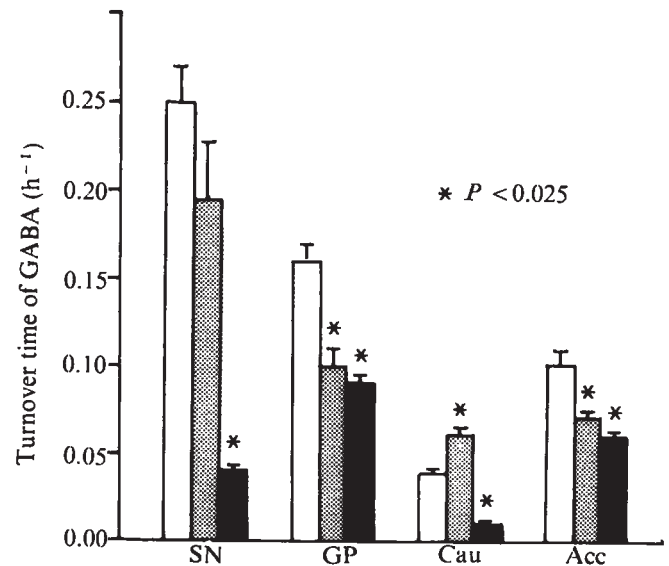


Fig. 1 Turnover time (h^{-1}) of GABA in various brain nuclei. Each bar represent an average of at least five measurements; brackets indicate s.e.m. Open bars, saline treated rats, closed bars, rats injected intraperitoneally with $30 \mu\text{mol kg}^{-1}$ of clozapine 20 min before ^{13}C glucose infusion. Dashed bars, rats treated intraperitoneally with $4 \mu\text{mol kg}^{-1}$ of haloperidol 40 min before labelling. SN, Substantia nigra, GP, globus pallidus; Cau, nucleus caudatus; Acc, nucleus accumbens.

Perhaps as proposed in earlier work¹⁶ an intrinsic positive feedback involving cholinergic synapses is operative; when this positive feedback was blocked by clozapine, the blockade of DA receptors failed to increase the ACh turnover rate in nucleus caudatus and nucleus accumbens.

Evidence is available to show that dopamine stored in dendrites of dopamine cell bodies of substantia nigra¹⁸ can be released by drugs¹⁹. In substantia nigra, the antipsychotics failed to change ACh turnover time (Table 1) suggesting that the blockade of DA receptors did not affect the ACh metabolism in substantia nigra (Table 1). However, cataleptogenic and non-cataleptogenic antipsychotics reduced the ACh metabolism in globus pallidus (Table 1).

Haloperidol and clozapine reduced the turnover time of GABA in globus pallidus and nucleus accumbens (Fig. 1). Both globus pallidus and nucleus accumbens contain high concentrations of GABA¹² and high GAD activity⁷. A pool of GABA-secreting interneurons, similar to that in nucleus caudatus⁴, may be present in nucleus accumbens and in globus pallidus²⁰. Since the blockade of DA receptors increased the rate of GABA metabolism in nucleus accumbens (Fig. 1), it is inferred that in this nucleus, DA receptors regulate GABA metabolism trans-synaptically. A similar suggestion can be made to explain the increase in GABA metabolism in globus pallidus because in this nucleus, which is devoid of dopaminergic terminals⁸, GABA may be stored in terminals from axons

Table 1 Effects of antipsychotics on acetylcholine (ACh) turnover time in various nuclei of rat brain

Brain nuclei	ACh turnover time (h^{-1})			
	Saline	Haloperidol ($10 \mu\text{mol kg}^{-1}$, i.p.)	Clozapine ($30 \mu\text{mol kg}^{-1}$, i.p.)	Chlorpromazine ($20 \mu\text{mol kg}^{-1}$, i.p.)
Substantia nigra (0.11 ± 0.018)	0.43 ± 0.07	0.31 ± 0.1	0.53 ± 0.06	0.63 ± 0.09
Globus pallidus (0.16 ± 0.023)	0.09 ± 0.03	$0.43 \pm 0.05^*$	$0.26 \pm 0.03^*$	$0.42 \pm 0.05^*$
Nucleus caudatus (0.60 ± 0.032)	0.10 ± 0.01	$0.04 \pm 0.01^*$	0.08 ± 0.01	$0.06 \pm 0.005^*$
Nucleus accumbens (0.60 ± 0.061)	0.10 ± 0.02	$0.04 \pm 0.01^*$	0.07 ± 0.01	$0.04 \pm 0.01^*$

Each value is the mean $\pm$ s.e.m. of at least five assays. Drugs were injected 51 min before the intravenous infusion of phosphoryl (d_4) choline ($15 \mu\text{mol kg}^{-1} \text{ min}^{-1}$) for 9 min. Numbers in parenthesis are the ACh concentration in $\mu\text{mol per g}$ protein (mean $\pm$ s.e.m.; $N = 5$). i.p., Intraperitoneal.

originating in nucleus accumbens. Haloperidol and clozapine act differently on ACh or GABA turnover in all structures tested except globus pallidus where they reduce ACh metabolism (Table 1) and increase that of GABA (Fig. 1). This increase of GABA utilisation in the nucleus accumbens-pallidal pathway decreases ACh metabolism in globus pallidus probably by a trans-synaptic mechanism.

Clozapine reduced GABA turnover time in substantia nigra and striatum, whereas haloperidol failed to change GABA metabolism in substantia nigra but decreased that of striatal GABA (Fig. 1). Since substantia nigra receives GABA terminals from nucleus caudatus²⁰⁻²³, the increase in the metabolism of GABA elicited by clozapine in substantia nigra and nucleus caudatus may be interrelated. At this time we can only suggest the possibility that muscarinic receptors regulate transmitter metabolism in intrinsic GABA interneurons of striatum and in the strio-nigra GABA neurones.

Indeed a precise interpretation for the mechanisms that are operative in contributing to the difference between the effects of haloperidol, chlorpromazine and clozapine on GABA and ACh metabolism in basal ganglia awaits further investigation; however, it is important to know that an enhanced metabolism of GABA in striatum and substantia nigra is elicited by clozapine, an antipsychotic which blocks DA receptors like haloperidol and chlorpromazine but unlike these two drugs, fails to cause tardive dyskinesias or extrapyramidal side effects^{24,25,26}.

In conclusion, the cataleptogenic antipsychotics decrease the ACh turnover time in nucleus caudatus and nucleus accumbens whereas clozapine does not; the latter decreases the turnover time of GABA in substantia nigra and nucleus caudatus whereas haloperidol does not. One might speculate on the possibility that antimuscarinic and GABA mimetic actions help to limit the cataleptogenic actions of antipsychotics; in contrast a reduction of ACh and an increase of GABA metabolism in globus pallidus express a characteristic change in striatal output elicited by cataleptogenic and non-cataleptogenic antipsychotics. Since globus pallidus includes an important output from nucleus caudatus and nucleus accumbens one might infer that the changes in ACh and GABA metabolism are biochemical responses related to the antipsychotic action which are independent from the extrapyramidal side effects.

E. MARCO
C. C. MAO
D. L. CHENEY
A. REVUELTA
E. COSTA

Laboratory of Preclinical Pharmacology,
National Institute of Mental Health,
Saint Elizabeths Hospital,
Washington, DC 20032

Received July 19; accepted September 27, 1976.

- 1 Anden, N. E., and Bedard, P., *J. Pharm. Pharmacol.*, **23**, 460-462 (1971).
- 2 Stadler, H., Lloyd, K. G., Gadea-Ciria, M., and Bartholini, G., *Brain Res.*, **55**, 476-480 (1973).
- 3 Trabucchi, M., Cheney, D. L., Racagni, G., and Costa, E., *Brain Res.*, **85**, 130-134 (1975).
- 4 McGeer, P. L., and McGeer, E. G., in *GABA in Nervous System Function*, (edit. by Roberts, E., Chase, C. N., and Tower, D. B.), 487-495 (Raven, New York, 1976).
- 5 Ungerstedt, U., *Acta physiol. scand. Suppl.*, **376**, 1-122 (1971).
- 6 Cheney, D. L., LeFevre, H. F., and Racagni, G., *Neuropharmacology*, **14**, 801-809 (1975).
- 7 Wood, J. G., McLaughline, B. J., and Vaugh, J. E., in *GABA in Nervous System Function*, (edit. by Roberts, E., Chase, C. N., and Tower, D. B.), 133-148 (Raven, New York, 1976).
- 8 Zivkovic, B., Guidotti, A., Revuelta, A., and Costa, E., *J. Pharmac. exp. Ther.*, **194**, 37-46 (1975).
- 9 Zsilla, G., Racagni, G., Cheney, D. L., and Costa, E., *Neuropharmacology*, (in the press).
- 10 Guidotti, A., Cheney, D. L., Trabucchi, M., Doteuchi, M., Wang, C., and Hawkins, R. A., *Neuropharmacology*, **13**, 1115-1122 (1974).
- 11 Zsilla, G., Cheney, D. L., Racagni, G., and Costa, E., *J. Pharmac. exp. Ther.*, (in the press).
- 12 Bertilsson, L., and Costa, E., *J. Chromat.*, **118**, 395-402 (1976).
- 13 Bertilsson, L., Mao, C. C., and Costa, E., *J. Pharmac. exp. Ther.*, (in the press).
- 14 Pearl, J., Spilker, A., Woodward, W. A., and Bentley, R. G., *J. Pharm. Pharmacol.*, **28**, 302-304 (1976).
- 15 Miller, R. J., and Hiley, C. R., *Nature*, **248**, 596-597 (1974).
- 16 Racagni, G., Cheney, D. L., Trabucchi, M., and Costa, E., *J. Pharmac. exp. Ther.*, **196**, 323-332 (1976).

- 17 Stadler, H., Gadea-Ciria, M., and Bartholini, G., *N.-S. Archs Pharmac.*, **288**, 1-6 (1975).
- 18 Bjorklund, A., and Lundvall, O., *Brain Res.*, **83**, 531-537 (1975).
- 19 Geffen, L. B., Jessel, T. M., Cuervo, A. C., and Iversen, L. L., *Nature*, **260**, 258-260 (1976).
- 20 McGeer, P. L., and McGeer, E. G., *Brain Res.*, **91**, 331-335 (1975).
- 21 Kim, J. S., Bak, I. J., Hassler, R., and Okada, Y., *Expl Brain Res.*, **14**, 95-104 (1971).
- 22 Yoshida, M., and Precht, W., *Brain Res.*, **32**, 225-228 (1971).
- 23 Precht, W., and Yoshida, M., *Brain Res.*, **32**, 229-233 (1971).
- 24 Crane, G. E., *Am. J. Psychiat.*, **124**, (Feb. Suppl.), 40-48 (1968).
- 25 Klawans, H. L., Jr., *Am. J. Psychiat.*, **130**, 82-86 (1973).
- 26 Shader, R. I., and DiMascio, A., *Psychotropic Drug Side Effects*, 92-106 (Williams and Wilkins, Baltimore, 1970).

Stimulation of pontine reticular formation suppresses firing of serotonergic neurones in the dorsal raphe

THERE is ample evidence¹⁻¹⁰ that both the pontine reticular formation (PRF; including the nucleus reticularis pontis oralis and caudalis in the terminology of Olszewski and Baxter¹¹ or the gigantocellular tegmental field in Berman's¹² terminology) and the brainstem raphe nuclei are involved in sleep processes. We report here that PRF-stimulation suppresses firing in raphe neurones. These results have implications which bear directly on current concepts about the cellular mechanisms involved in the control of the sleep cycle.

Evidence from lesion¹⁻³ and stimulation⁴⁻⁸ studies suggests that the PRF is important in the generation of rapid eye movement (REM) or desynchronised sleep (DS). Hobson *et al.*⁶⁻⁸ have found that cells in the PRF selectively increased their firing just before the onset of DS and maintained their high activity throughout the whole DS episode. On the other hand, McGinty and Harper⁹ reported that cells in the dorsal raphe nucleus (DRN) ceased their firing before the occurrence of pontogeniculooccipital (PGO) waves and their activity remained low during the ensuing DS. The occurrence of PGO waves is thought to be inhibited or 'gated' by the release of 5-hydroxytryptamine (5-HT)^{13,14}. It is possible that immediately preceding the onset of and during the DS episode, cells in the PRF exert an inhibitory influence on cells containing 5-HT in the DRN which, in turn, release the gate or disinhibit the PGO wave activity. Anatomical studies using Golgi¹⁵, Nauta¹⁶, autoradiography¹⁷ and horseradish peroxidase¹⁸ techniques have indicated the presence of projections from the PRF to midbrain raphe nuclei. We have also presented pharmacological evidence¹⁹ suggesting the existence of an inhibitory pathway from the PRF to the DRN. The purpose of this study was to investigate the influence of electrical stimulation of the PRF on 5-HT cells in the DRN and to explore the possible neurotransmitters which might mediate these effects.

Under chloral hydrate anaesthesia (400 mg kg⁻¹, intraperitoneally), 20 male Sprague-Dawley rats (200-300 g) were implanted with concentric stimulating electrodes in at least one, in most cases two, of the following structures: (1) the ventromedial PRF [posterior 1 mm (to lambda); lateral 0.7 mm; horizontal -8 mm (below skull surface)]; (2) the locus coeruleus (LC; P 1 mm; L 1.1 mm; H -6 mm); and (3) the nucleus trigemini principalis (NprV; P 1 mm; L2 to 2.5 mm; H -6.5 mm). Stimulation of the latter two structures served as controls. Electrodes were positioned at an angle of 10° to the vertical so that there was enough space for lowering the recording micropipettes. Micropipettes filled with 2 M NaCl saturated with fast green FCF²⁰ were lowered through a burr hole into the DRN (A 0.5 mm; L 0 mm; H -5.5 to -7 mm). Methods for differential recording of single unit activity during electrical stimulation, drug administration, histological localisation and physiological data analysis have already been described in detail^{19,20}. 5-HT cells in the DRN were tentatively identified by their wave form (that is, a predominant positive then a negative wave) and by their slow, regular spontaneous dis-

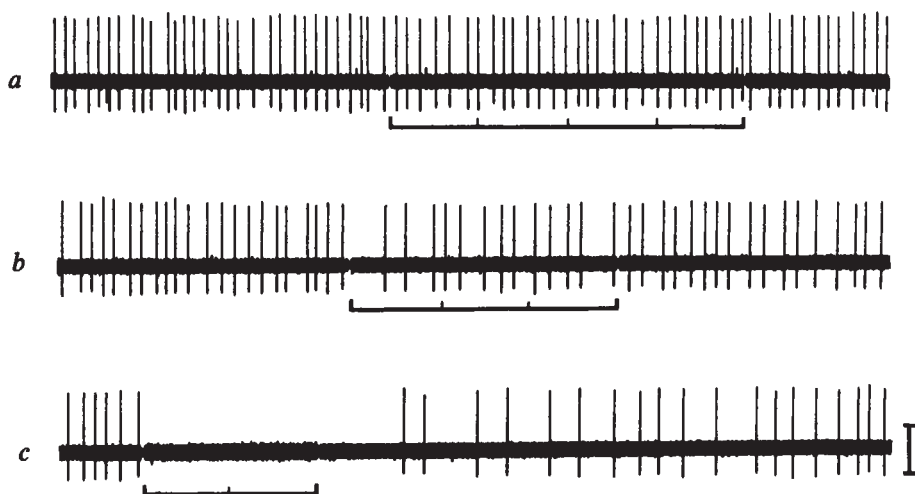


Fig. 1 Comparison between the responses of a typical 5-HT cell to PRF-stimulation and control electrode stimulation. This experiment was carried out in a 6-OHDA pre-treated (8 d) rat to rule out the involvement of the LC and other adrenergic systems. We gave an intraperitoneal injection of chlorimipramine (1 mg kg^{-1}) to protect the 5-HT system before the intraventricular injection of $200 \mu\text{g}$ of 6-OHDA in $20 \mu\text{l}$ of ascorbic acid solution (1 mg ml^{-1}). *a-c*, Continuous oscillographic records of cellular discharge. The line under the oscillographic record indicates the stimulation period. Each division on the line is 10 s. In *a*, there was no suppression following the stimulation of the control electrode (10 Hz, 0.1 mA, 1 ms pulsewidth) which was located in the NprV. In *b*, PRF-stimulation at 0.05 mA (10 Hz, 1 ms) resulted in a slowing of the firing. In *c*, PRF-stimulation at 0.1 mA (10 Hz, 1 ms) produced a total suppression of cellular activity. Calibration: 0.5 mV. Positivity is upward.

charge rates ($0.5\text{--}2.5 \text{ spikes s}^{-1}$). This firing pattern has been demonstrated by combined single cell recording and fluorescence histochemical methods²² to be characteristic for 5-HT-containing cells but not for non-5-HT-containing cells in the adjacent central grey or reticular formation. Recently, the identification of the slow, regular units in the DRN as 5-HT neurones has been confirmed electrophysiologically: electrical stimulation of the ascending 5-HT axonal pathway in the ventromedial tegmentum of the anterior midbrain produces antidromic responses only in the type of unit described above (ref. 21 and R.Y.W. and G.K.A., in preparation).



Fig. 2 Photomicrograph of a coronal section of the brain stained with cresyl violet showing that in a bilateral LC(*)-lesioned rat, the tract of the PRF stimulating electrode (arrow) is in the nucleus reticularis pontis caudalis.

Electrical stimulation of the ventromedial PRF markedly suppressed spontaneous firing of 5-HT cells in the DRN; in many cases suppression of firing was produced by PRF-stimulation at relatively low frequencies and currents. The term 'suppression' is used in preference to 'inhibition' because we cannot rule out the possibility that the depressant effects of PRF-stimulation on DRN cells are due (indirectly) to disfacilitation. The degree of suppression of the firing of DRN cells was proportional to the strength of the stimulus applied to the PRF (Fig. 1*b* and *c*). Repeated stimulation at 1 Hz ($0.05\text{--}0.2 \text{ mA}$, pulsewidth 1 ms) resulted in a poststimulus period of total suppression lasting 50–300 ms in 66 of 68 (97%) typical 5-HT cells. Of these 66 cells, 7 showed an initial excitation preceding the period of suppression. None of the cells ($N = 4$) which lacked the characteristic firing pattern of 5-HT neurones was affected by the stimulation. Most of the 5-HT cells responded to the PRF-stimulation with a short latency of suppression (range, 0–66 ms; mode 0 ms; median, 8 ms; mean $\pm$ s.e., $10.0 \pm 11.5 \text{ ms}$). By 0 latency ($N = 14$) we mean that the onset of the suppression was so rapid that no spike occurred in between onsets of the stimulation and the suppression of firing. Although the short latency of this effect suggests the existence of a monosynaptic PRF–DRN pathway, a polysynaptic pathway cannot be ruled out by our data. The mechanism of PRF-induced suppression of firing cannot be attributed to recurrent inhibition because none of the recorded raphe cells showed signs of antidromic activation (that is, invariant latency, faithful response to high rate of stimulation and collision to orthodromic spikes). The absence of antidromic responses might be due to the fact that there are only scattered 5-HT terminals in the reticular formation²³.

In four experiments, stimulating electrodes were put into different regions of the anterior medullary reticular formation. In these rats, 5-HT cells ($N = 6$) in the DRN were not suppressed by the stimulation. This suggests that the effects of PRF stimulation are mediated by cells which originate in this

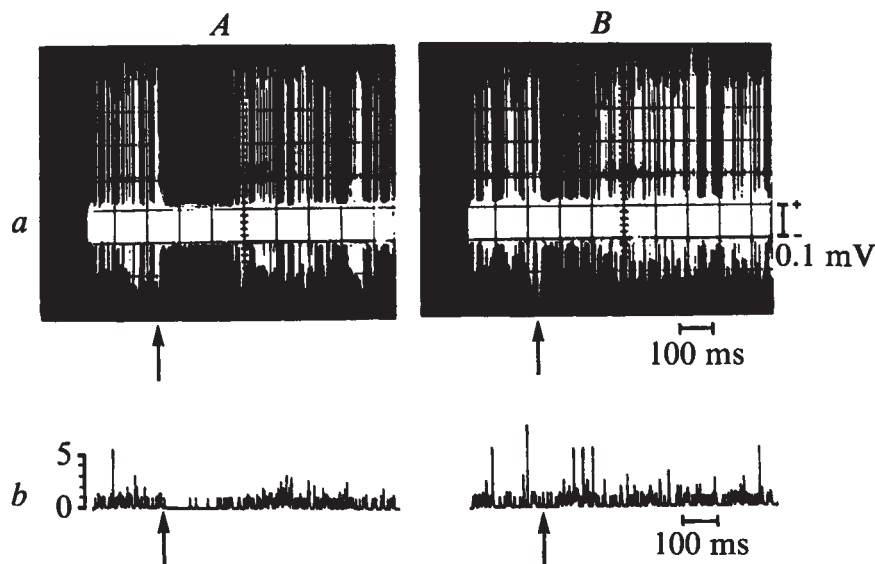
Table 1 Comparison between the response of 5-HT cells to PRF stimulation and to control electrode stimulation

Electrode location	No. of rats	No. of cells suppressed	No. of cells not suppressed	% Suppressed
PRF	18	66	2	97%
LC*	4	4	15	21%
NprV*	9	3	19	14%

PRF, pontine reticular formation; LC, locus coeruleus; NprV, nucleus trigemini principalis.

*Compared to PRF stimulation, the proportion of 5-HT cells suppressed by stimulation of control electrodes is significantly lower ($P < 0.001$, χ^2 test with Yate's correction). The difference between the LC and the NprV groups is not significant.

Fig. 3 PRF-induced suppression of firing of 5-HT neurones and their blockade by intravenous injection of the GABA antagonist picrotoxin. *a* and *b* are two different 5-HT cells from two rats in separate experiments. Activity of neurones was displayed either directly on a storage oscilloscope (cell *a*) or in the form of a peri-stimulus time histogram (cell *b*). Every picture and histogram consists of at least 150 sweeps. The latency of suppression was determined by measuring the time elapsed between the onset of stimulation and the last spike preceding the onset of suppression. Arrows indicate the onset of electrical stimulation. In column *A*, before administration of picrotoxin, PRF-stimulation had a marked effect on 5-HT neurones. In column *B*, after the intravenous injection of picrotoxin, PRF-induced suppression was reduced in cell *a* (2.5 mg kg^{-1}) and was totally blocked in cell *b* (6 mg kg^{-1}). Vertical counts in cell *b* represent number of spikes in each address.



area rather than ascending fibres. This agrees with the anatomical study of Nauta¹⁶ who has shown that the segment of the ascending component of Forel's tractus fasciculorum which projects to the midbrain raphe area originates at a more rostral tegmental level rather than from the medulla.

When the stimulating electrode was located either in the LC or in the NprV, there were considerably fewer 5-HT cells suppressed by the stimulation. Table 1 lists the number of 5-HT cells suppressed by stimulation of the PRF, LC and NprV. Compared to PRF-stimulation, the proportion of 5-HT cells suppressed by stimulation of either the LC or the NprV is significantly smaller ($P < 0.001$, χ^2 test); latencies of suppression produced by stimulation of control electrodes (combined LC and NprV group) were longer ($16.5 \pm 15.9 \text{ ms}$ against $10.0 \pm 11.5 \text{ ms}$, $P > 0.05$, two-tailed t test); and current thresholds for control electrodes to elicit suppression of firing on 5-HT cells were much higher ($0.27 \pm 0.06 \text{ mA}$ against $0.09 \pm 0.04 \text{ mA}$, $P < 0.001$, two-tailed t test; Fig. 1*a*).

To demonstrate further that stimulation of the LC does not contribute materially to PRF-induced suppression, we bilaterally destroyed the LC before the stimulation and recording experiments. Using either radio-frequency lesions or chemical lesions with intraventricular injection of 6-hydroxydopamine (6-OHDA), we have consistently destroyed the LC¹⁹. The extent of the LC lesions was verified by either routine histological methods (Fig. 2) or the Falck-Hillarp type fluorescence histochemical method. As can be seen from Fig. 2, the radio frequency lesion destroyed the dorsolateral part of the PRF as well as the LC. In these LC (and dorsolateral PRF)-lesioned rats ($N = 6$) PRF-induced suppression of 5-HT cells was undiminished; all 15 5-HT cells were readily inhibited by ventromedial PRF-stimulation (Fig. 1*b* and *c*).

In an attempt to explore the possible transmitters involved in the PRF-induced suppression, we compared the PRF-induced effect on 5-HT cells both before and after the intravenous administration of the γ -aminobutyric acid (GABA) antagonist picrotoxin or the glycine antagonist strychnine. We found that picrotoxin reduced markedly the PRF-induced suppression of 5-HT cells ($N = 6$). Figure 3 is a representative example showing that picrotoxin reduced PRF-induced suppression at a dose of 2.5 mg kg^{-1} and totally blocked the suppression at a dose of 6 mg kg^{-1} . In contrast, picrotoxin at these dose levels, failed to block the recurrent inhibition of DRN cells produced by stimulation of the ascending 5-HT pathway (R.Y.W. and G.K.A., in preparation). Strychnine ($0.4\text{--}1 \text{ mg kg}^{-1}$, intravenously) had little or no effect on PRF-induced suppression (8 cells). These results are suggestive of a possible role for GABA in mediating the effect produced by PRF-stimulation and are consistent with our previous finding

that 5-HT cells in the DRN are inhibited by local, iontophoretic application of GABA and that this inhibition is blocked by picrotoxin¹⁹.

In conclusion, our data indicate that stimulation of the PRF markedly suppresses the activity of 5-HT cells in the DRN and that this effect might be mediated through a PRF-DRN GABAergic pathway. These results support the hypothesis that immediately preceding the onset of and continuing throughout the DS episode, cells in the PRF, which are known to increase their firing at this time⁶⁻⁸, exert a powerful, inhibitory influence on 5-HT cells in the DRN. On the other hand, we conclude that the LC has a minor and possibly indirect influence on 5-HT cells in the DRN because: (1) stimulation of the LC (which consists of cells containing noradrenaline) does not have a marked inhibitory effect on 5-HT cells in the DRN; and (2) the microiontophoretic application of noradrenaline directly on to 5-HT cells does not consistently produce inhibition^{19,24}. Experiments are now being conducted on behaving rats to eliminate the possible influence of anaesthesia on the response of DRN cells to PRF stimulation. *Note added in proof:* A recent abstract²⁵ agrees with our result that stimulation of LC has a minor and possibly indirect influence on 5-HT cells in the DRN.

R. Y. WANG
D. W. GALLAGER*
G. K. AGHAJANIAN

Departments of Psychiatry and Pharmacology,
Yale University School of Medicine and
the Connecticut Mental Health Center,
34 Park Street, New Haven, Connecticut 06508

Received July 19; accepted September 27, 1976.

*Present address: NIMH, NIH Bethesda, Maryland 20014.

- Jouvet, M., *Arch. Ital. Biol.*, **100**, 125-206 (1962).
- Carli, G., and Zanchetti, A., *Arch. Ital. Biol.*, **103**, 751-788 (1965).
- Hobson, J. A., *Electroencephalog. clin. Neurophysiol.*, **19**, 41-62 (1965).
- Monti, J. M., *Expl Neurol.*, **28**, 484-493 (1970).
- Frederickson, C. J., and Hobson, J. A., *Arch. Ital. Biol.*, **108**, 564-576 (1970).
- Hobson, J. A., McCarley, R. W., Pivik, R. T., and Freedman, R., *J. Neurophysiol.*, **37**, 497-511 (1974).
- Hobson, J. A., McCarley, R. W., Freedman, R., and Pivik, R. T., *J. Neurophysiol.*, **37**, 1297-1309 (1974).
- Hobson, J. A., McCarley, R. W., and Wyzinski, P. W., *Science*, **189**, 55-58 (1975).
- McGinty, D. J., and Harper, R. M., *Brain Res.*, **101**, 569-575 (1976).
- Jouvet, M., *Science*, **163**, 32-41 (1969).
- Olzsewski, J., and Baxter, D., *Cytoarchitecture of the Human Brain Stem* (Karger, New York and Basel, 1954).
- Berman, A. L., *The Brain Stem of the Cat* (University of Wisconsin Press, Madison, 1968).
- Brooks, D. C., and Gershon, M. D., *Brain Res.*, **27**, 223-239 (1971).
- Simon, R. P., Gershon, M. D., and Brooks, D. C., *Brain Res.*, **58**, 313-330 (1973).
- Scheibel, M. E., and Scheibel, A. B., in *Reticular Formation of the Brain* (edit. by Jasper, H. H., Proctor, L. D., Kington, R. S., Noshay, R. T., and Costello, R. T.), 31-55 (Little Brown, Boston, 1958).
- Nauta, W. J. H., and Kuypers, H. G. J. M., *ibid.*, 3-30.
- Büttner-Ennever, J. A., and Henn, V., *Brain Res.*, **108**, 155-164 (1976).

- 18 Aghajanian, G. K., and Wang, R. Y., *Brain Res.* (in the press).
 19 Gallager, D. W., and Aghajanian, G. K., *Eur. J. Pharmac.* (in the press).
 20 Thomas, R. C., and Wilson, V. J., *Nature*, **206**, 211–213 (1965).
 21 Wang, R. Y., and Aghajanian, G. K., *Brain Res.* (in the press).
 22 Aghajanian, G. K., and Haigler, H. J., *Brain Res.*, **81**, 364–372 (1974).
 23 Aghajanian, G. K., Haigler, H. J., and Bennett, J. L., in *Handbook of Psychopharmacology*, 6 (edit. by Iversen, L. L., Iversen, S. D., and Snyder, S. H.), 63–96 (Plenum, New York, 1975).
 24 Svensson, T. H., Bunney, B. S., and Aghajanian, G. K., *Brain Res.*, **92**, 291–306 (1975).
 25 Anderson, et al., *Neurosci. Abstr.*, **2**, 477 (1976).

Decrease of acetylcholine receptor synthesis in muscle cultures by electrical stimulation

MUSCLE activity has been shown to be a central factor in controlling the level of extrajunctional acetylcholine receptors (AChR): inactivity of muscle fibres causes an increase in the number of AChR and electrically stimulated activity causes a decrease. We investigated the mechanism by which this activity regulates the level of AChR in muscle fibres that have differentiated in cell culture. Our results suggest that electrical stimulation decreases the AChR level by decreasing the synthesis of the receptors rather than by degrading or inactivating them.

In normally innervated muscle, AChR are located at the endplate. After surgical denervation, however, there is a great increase in the number of AChR in the region outside the endplate¹. A clear indication that muscle activity plays a part in controlling the extrajunctional AChR came from the experiments in which an electrically stimulated contraction of denervated muscle suppressed the spread of

Table 1 Effect of cycloheximide on AChR

	AChR (fmol per dish)	Half life (h)
Control	219 ± 20	28 ± 3
Cycloheximide (20 µg ml ⁻¹)	231 ± 16	267 ± 22
Stimulated	123 ± 4	26 ± 2
Stimulated + cycloheximide	224 ± 14	276 ± 26

10-d-old chick muscle cultures were treated for 48 h as indicated, then the amount of toxin bound, which reflects the amount of AChR, was measured. To measure the rate of AChR degradation, 14-d-old muscle cells were first pulsed with ¹²⁵I-α-BuTX then the cells were treated as indicated for 3 d. Half life of AChR was determined from the decline of radioactivity bound in the culture dish. Each value is the average of three replicates ± s.d.

AChR^{2,3}. Furthermore, it was reported^{4–7} that electrical stimulation of denervated muscle reduced the supersensitivity to ACh.

The reduction in the amount of AChR following electrical stimulation could be a result of: (1) inhibition of AChR synthesis, (2) increase in AChR degradation, (3) inactivation of extrajunctional AChR. To distinguish between these possibilities we used muscle cells grown *in vitro*. Such muscle fibres were found to produce AChR⁸ and also to develop ACh supersensitivity in response to tetrodotoxin (TTX) application, which inhibits their spontaneous activity^{9,10}. When muscle cultures were stimulated electrically, the amount of AChR, as measured by labelled α-bungarotoxin (α-BuTX) binding, was reduced considerably, compared to the control or to the TTX-treated cells (Fig. 1).

The above-mentioned difference between stimulated cultures and those of the control could be explained by a different turnover rate of AChR. Such a difference in the turnover rate of junctional and extra-junctional receptors has been reported^{11,12}. To test this possibility, muscle cultures were first pulsed with ¹²⁵I-α-BuTX, then electrically stimulated and the rate of toxin loss was measured (Table 1). If we assume that the toxin loss reflects intracellular degradation of toxin-receptor complex¹³, then the results show that the various treatments had very little effect on AChR degradation rate. (The half life fell within a 26–28 h range.) When protein synthesis was inhibited in the electrically stimulated cultures by cycloheximide, no decrease in the amount of AChR was observed (Table 1). This result is in contrast to the above-mentioned decrease in AChR resulting from electrical stimulation in the absence of protein synthesis inhibitor. This effect of cycloheximide might be explained, as has been suggested^{11,13}, by an inhibition of receptor degradation (Table 1). The fact that there is no decrease in the amount of AChR in the presence of the inhibitor suggests that the electrical stimulation *per se* does not inactivate the receptors.

To measure the effect of electrical stimulation on receptor synthesis, the existing cholinergic receptors were saturated with unlabelled α-BuTX. Then the rate of appearance of new receptors was measured by binding of radioactive α-BuTX after various periods of stimulation. It was reproducibly found that the stimulated cultures bound much less ¹²⁵I-α-BuTX than did the control, or the TTX-treated cultures (Fig. 2). These experiments suggest that the decrease in α-BuTX binding as a consequence of electrical stimulation is due to a partial suppression in receptor synthesis rather than an inactivation of existing AChR, or an increase in the rate of their turnover.

Walker and Wilson¹⁴ have reported that the electrical stimulation of cultured muscle cells caused a 30–90% decrease in acetylcholinesterase (AChE) activity. Although we used the same technique for measuring the enzyme activity, our results differed considerably; in our case the decrease in AChR after electrical stimulation was accompanied by no significant change in AChE activity. Since cell injury might be the cause of the decreased AChE activity,

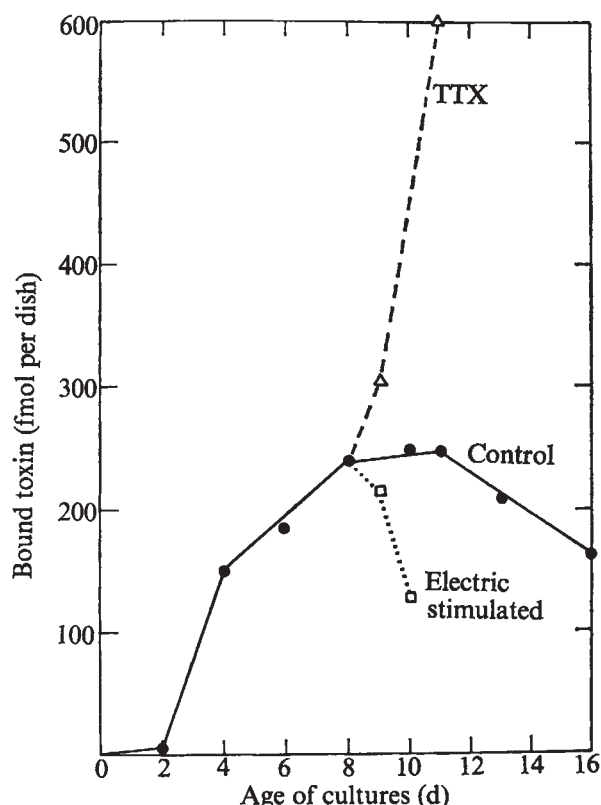


Fig. 1 Time course of ¹²⁵I-α-BuTX binding by chick muscle cultures treated with electrical stimulation (□), or 0.5 µg ml⁻¹ TTX (△). Each point represents the mean of three replicates. The electrical stimulation of the muscle was achieved through a pair of platinum electrodes immersed in the culture medium and fastened 25 mm apart, to a cover glass of the 30 mm dish. Six-second trains of 1 ms square-wave pulses with an interval between pulses of 200 ms were delivered to each dish every minute. Pulse polarity was alternated for each train to prevent polarisation. The voltage was approximately 20 V. Growing of the muscle cells and the determination of the amount of toxin bound were as described before¹⁰.

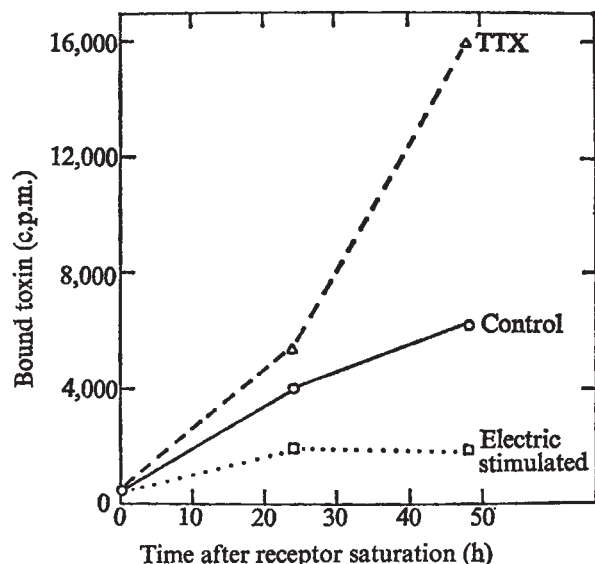


Fig. 2 Kinetics of ^{125}I - α -BuTX binding after saturation of AChR with cold toxin. The appearance of new receptors in 9-d-old muscle cultures was followed with labelled α -BuTX during treatment with electrical stimulation ($\square$), or TTX (Δ).

we monitored possible cell damage in our experiments by measuring the activity of creatine phosphokinase (CPK), which was established as an indicator for myotube proteins¹⁵. In the experimental conditions which we chose, the electrical stimulation caused no change in CPK activity, nor was any cell damage observed (such damage was observed if the stimulation was given at a much higher frequency). Furthermore, since total protein synthesis and nucleic acid synthesis in stimulated cultures were similar to those in controls, we assume that the effect of electrical stimulation on synthesis was not a general one, but specifically affected the receptor synthesis machinery.

In conclusion, muscle fibres differentiated *in vitro* in the absence of nervous tissue have been used as a model system for the study of molecular mechanism accompanying denervation and reinnervation without the effect of a neurotrophic factor. While inactivity, like denervation, induced receptor synthesis, mechanical activity suppressed it. These results do not exclude some additional trophic influence of nerve on the muscle along with the effect of muscle usage. The existence of such an additional neurotrophic effect has recently received further support^{16,17}.

This work was supported by grants from the Muscular Dystrophy Association of America, and the US-Israel Binational Science Foundation, Jerusalem, Israel.

ASHER SHAINBERG
MATI BURSTEIN

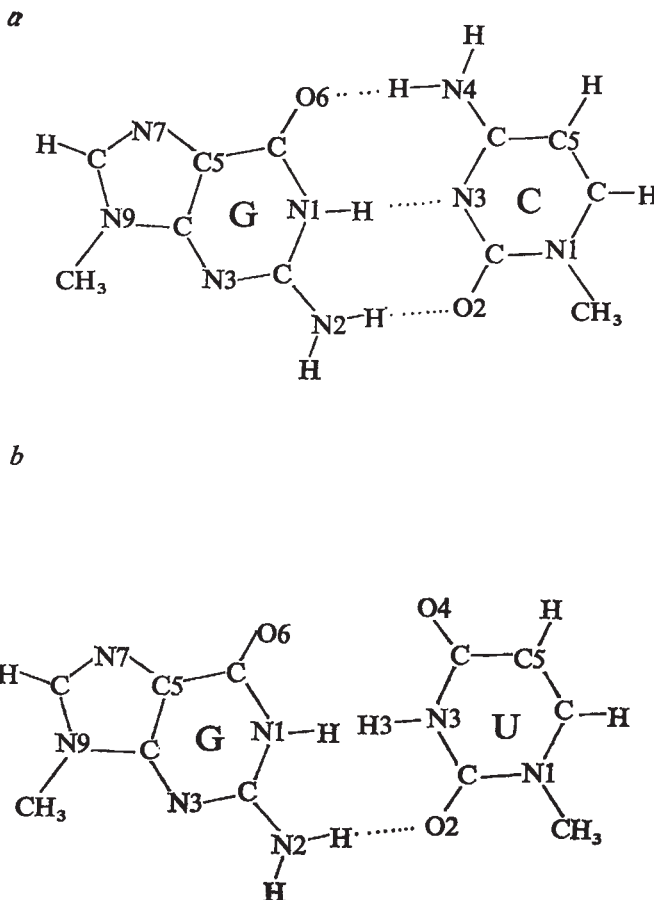
Department of Life Sciences,
Bar-Ilan University,
Ramat-Gan, Israel

Effect of guanine alkylation on mispairing

ALKYLATING agents, such as nitrosamines, nitrosoureas, dimethylsulphate, and so on, are known to cause mutations and cancerous growth¹⁻³. Since nitrosamines can be produced by the reaction of amines and nitrous acid, the presence of nitrites and nitrous acid in many food items as well as in the urban environment has become a source of concern^{4,5}. Recent evidence⁶ has indicated that carcinogenesis and mutagenesis bear a strong relationship to each other, although agreement is not unanimous⁷. *In vivo* alkylation of DNA occurs mainly at the N7 position of guanine (G)¹⁻². Minor alkylation of G at the O6 and N3 positions is also known to occur⁸⁻⁹. Whether alkylation at any one of these sites is mainly responsible for mutagenesis and carcinogenesis is still an open question, although O6 alkylation seems to be favoured^{3,8-11}. We have investigated the effect of methylation at each of these three sites on the ability of G to mispair with a thymine (or uracil), which may lead to a mutation, and the mechanism involved in such mispairing, using the CNDO/2 method of calculation. The calculations are carried out for the 9-methyl purines and 1-methyl pyrimidines to simulate the bases in the DNA helix. The details of the calculations are published elsewhere¹². The magnitudes reported are particular to the geometrical parameters used here¹³. Since we were only interested in the relative effect of various G methylations, the qualitative trends observed here are expected to hold; no attempt at a full minimisation of the geometric parameters was therefore made; (a horrendous task for molecules of such sizes).

Figure 1 shows the normal guanine-cytosine (G-C) and the

Fig. 1 G-C (a) and G-U base pairs (b).



Received July 16; accepted October 4, 1976.

- 1 Axelsson, J., and Thesslef, S., *J. Physiol., Lond.*, **147**, 178-193 (1959).
- 2 Lomo, T., and Rosenthal, J., *J. Physiol., Lond.*, **221**, 493-513 (1972).
- 3 Drachman, D. B., and Witzke, F., *Science*, **176**, 514-516 (1972).
- 4 Purves, D., and Sakmann, B., *J. Physiol., Lond.*, **237**, 157-182 (1974).
- 5 Hogan, P. G., Marshall, J. M., and Hall, Z. W., *Nature*, **261**, 328-330 (1976).
- 6 Jones, R., and Verbova, G., *J. Physiol., Lond.*, **236**, 517-538 (1974).
- 7 Lomo, T., and Westgaard, R. H., *J. Physiol., Lond.*, **252**, 603-626 (1975).
- 8 Vogel, Z., Sytkowski, A. J., and Nirenberg, M. W., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 3180-3184 (1972).
- 9 Cohen, S. A., Fischbach, G. D., *Science*, **181**, 76-78 (1973).
- 10 Shainberg, A., Cohen, S. A., and Nelson, P. G., *Pflügers Arch. Eur. J. Physiol.*, **361**, 255-261 (1976).
- 11 Berg, D. K., and Hall, Z. W., *J. Physiol., Lond.*, **252**, 779-790 (1975).
- 12 Chang, C. C., and Huang, M. C., *Nature*, **253**, 643-644 (1975).
- 13 Devreotes, P. N., and Fambrough, D. M., *J. Cell Biol.*, **65**, 335-358 (1975).
- 14 Walker, C. R., and Wilson, B. W., *Nature*, **256**, 215-216 (1975).
- 15 Shainberg, A., Yagil, G., and Yaffe, D., *Dev. Biol.*, **25**, 1-29 (1971).
- 16 Lavoie, P. A., Collier, B., and Tenenhouse, A., *Nature*, **260**, 349-350 (1976).
- 17 Pestronk, A., Drachman, D. B., and Griffin, J. W., *Nature*, **260**, 352-353 (1976).

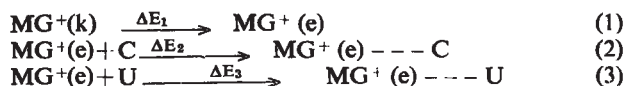
Table 1 Interaction energies (kcalorie mol⁻¹) of G and M G for the processes in mechanisms a (upper line) and c (bracketed values in lower line)

	G	N7 MG	O6 MG	N3 MeG
ΔE_1	12.70 (513.91)	13.76 (408.65)	100.33 (371.49)	-2.37 (375.90)
ΔE_2	22.06 (-12.83)	16.37 (-13.53)	14.42 (-7.03)	11.60 (-16.99)
$\Delta E_3 - \Delta E_2$	-37.34 (7.84)	-37.05 (7.77)	-44.39 (-0.70)	-33.49 (10.95)
$\Delta E_T - \Delta E_T^*$	0 (0)	1.35 (-105.33)	80.58 (-150.96)	-11.22 (-134.90)

* $\Delta E_T = (\Delta E_3 - \Delta E_2) + \Delta E_1$; $\Delta E_T^* = \Delta E_T$ for unmethylated G.

anomalous guanine-uracil (G-U) base pairs. Whereas the former base pair is stable in the present calculations by -22.64 kcalorie mol⁻¹, the latter is unstable by 20.03 kcalorie mol⁻¹. The stability of G-U can be enhanced by increasing the number of hydrogen bonds between G and U. This may be achieved by the interaction of: (a) a-G in the enol (e) form, where the H on N₁ is moved to the O6 position, with U in the keto (k) form; (b) an-U in the e form, where the H on N3 is shifted to position O4, with G(k). (We have already shown that this type of mispairing takes place in the case of 5-fluorouracil¹², a known mutagen); (c) an ionised G, in which the H on N1 is removed, with U(k). This mechanism was proposed by Lawley and Brookes¹⁴, who found a large increase in the extent of such ionisation on alkylation at the N7 position. Mechanisms a and b follow the original explanation for mispairing proposed by Watson and Crick¹⁵.

In mechanism a one must estimate the energies for each of the three equations:



where MG⁺ indicates a methylated G cation. We will now denote the specific site of alkylation by a subscript to M. Table 1 shows that the e form is favoured over the k form only in the case of M₃G⁺, and that alkylation of the N3 position is the only one that leads to a greater probability for mispairing by this mechanism than G itself.

Mechanism b is very improbable due to the very low probability of finding a U in the e form¹². Nevertheless, our calculations for the competitive interaction of a G(or MG⁺) with a C(k) compared with a U(e), predict that most mispairing would occur in the case of the unalkylated G. This mechanism therefore cannot explain the increase in mutation on alkylation.

In mechanism c one must compare the energies for the following steps involving the various purine species:



These energy values are shown in brackets in Table 1. The ΔE_1 values are obtained as the difference between the energies of MG(k) and MG⁺(k) for the methylated species and of G⁺(k) and G(k) for the unalkylated G. To obtain reliable estimates of the overall energy change in solution for step 1, one must tackle the difficult problem of solvent interaction^{16,17} with each of the purine species as well as with the H⁺ involved in the ionisation of the purines. Nevertheless, the predicted qualitative trend of an increase in ionisation on alkylation at N7 is consistent with experimental observation¹⁴. The relative differences between the ΔE_1 values for the methylated species should still

hold upon inclusion of solvent effects because of the expectedly quite close MG⁺-solvent interaction energies of the three methylated cations. Note that the greatest extent of ionisation is predicted for M₆G⁺, and that on losing the proton it is the only species that prefers to pair with a U(k) rather than a C(k), reflected in its negative ($\Delta E_3 - \Delta E_2$) value. The total relative energies in the last row of Table 1 shows that all three types of methylation increased the extent of mispairing by this mechanism in the order: M₆G⁺ > M₃G⁺ > M₇G⁺.

The results of mechanism c are consistent with the experimental ones of Gerchman and Ludlum¹¹. These authors found that O6 methylated G mispairs to U much more than N7 methylated G or G itself; and that this greater extent of mispairing was unaffected by the presence of C as a competitor.

Our results predict that methylation at the O6 and N3 positions is more effective than at the N7 position in leading to mispaired bases and therefore to mutations and possible cancerous growth. Whereas mispairing occurs by mechanism c for O6—and N7—methylated G, for N3 methylated G it may occur by either or both mechanisms a and c. The relative importance of these two mechanisms must await experiments on the relative extent of ionisation compared with enol formation for M₃G⁺.

Our calculations indicate that methylation at the O6, N7 and N3 positions of G increases the stability of the G-C base pair by -3.80, -5.22 and -11.54 kcalorie mol⁻¹, respectively. If all other interactions within the DNA helix remain unaffected (a most unlikely situation) the methylation of G while in the helical structure would therefore help stabilise the helix against denaturation. The resulting reduction in DNA replication may be responsible for the somewhat weak chemotherapeutic action of monoalkylating agents¹⁸.

We thank the NSF for financial support.

S. F. ABDULNUR
R. L. FLURRY, JR

Department of Chemistry,
University of New Orleans,
Lakefront, New Orleans, Louisiana 70122

Received July 15; accepted September 27, 1976.

- 1 Magee, P. N., and Barnes, J. M., *Adv. Cancer Res.*, **10**, 163-246 (1967).
- 2 Lawley, P. D., *Progr. Nucl. Acid Res. molec. Biol.*, **5**, 89-133 (1966).
- 3 Magee, P. N., Nicoll, J. W., Pegg, A. E., and Swann, P. G., *Biochem. Soc. Trans.*, **3**, 62-65 (1975).
- 4 Shapley, D., *Science*, **191**, 268-270 (1976).
- 5 Wolff, I. A., and Wasserman, A. E., *Science*, **177**, 15-19 (1972).
- 6 Ames, B. N., *Science*, **191**, 241-245 (1976).
- 7 Rubin, H., *Science*, **191**, 241 (1976).
- 8 Loveless, A., *Nature*, **223**, 206-207 (1969).
- 9 Lawley, P. D., Orr, D. J., and Shah, S. A., *Chem. Biol. Int.*, **4**, 431-434 (1972).
- 10 Lawley, P. D., and Shah, S. A., *Chem. Biol. Int.*, **5**, 286-288 (1972).
- 11 Gerchman, L. L., and Ludlum, D. B., *Biochim. biophys. Acta*, **308**, 310-316 (1973).
- 12 Abdunur, S. F., *J. theor. Biol.*, **58**, 165-175 (1976).
- 13 Breen, D. L., and Flurry, R. L., Jr., *Theor. Chim. Acta., Berl.*, **23**, 138-146 (1971).
- 14 Lawley, P. D., and Brookes, P., *Nature*, **192**, 1081-1082 (1961).
- 15 Watson, J. D., and Crick, F. H. C., *Cold Spring Harb. Symp. quant. Biol.*, **18**, 123-131 (1953).
- 16 Abdunur, S. F., thesis, Yale Univ. (1966), Dissertation Abstracts, 4271-B (1967).
- 17 Sinanoglu, O., and Abdunur, S. F., *Fedn Proc.*, **24**, S12-S23 (1965).
- 18 Lawley, P. D., Lethbridge, J. H., Edwards, P. A., and Shooter, K. V., *J. molec. Biol.*, **39**, 181-198 (1969) and refs therein.

Interferon, double-stranded RNA and mRNA degradation

INTERFERONS are glycoproteins which are synthesised in various animal cells following viral infection. They are excreted, interact with other cells and inhibit in them the replication of a broad range of viruses¹. Extracts from interferon-treated Ehrlich ascites tumour cells (S30_{INT}) differ in various biochemical characteristics from extracts from control cells (S30_C) (refs 2-4 and see also refs 5-8). We have reported recently that reovirus messenger RNAs (mRNAs) are degraded faster in reaction mixtures containing S30_{INT} than in those containing S30_C, but only if the reaction mixtures are supplemented with double-stranded (ds) RNA⁹. We have also reported¹⁰ that dsRNA promotes the phos-

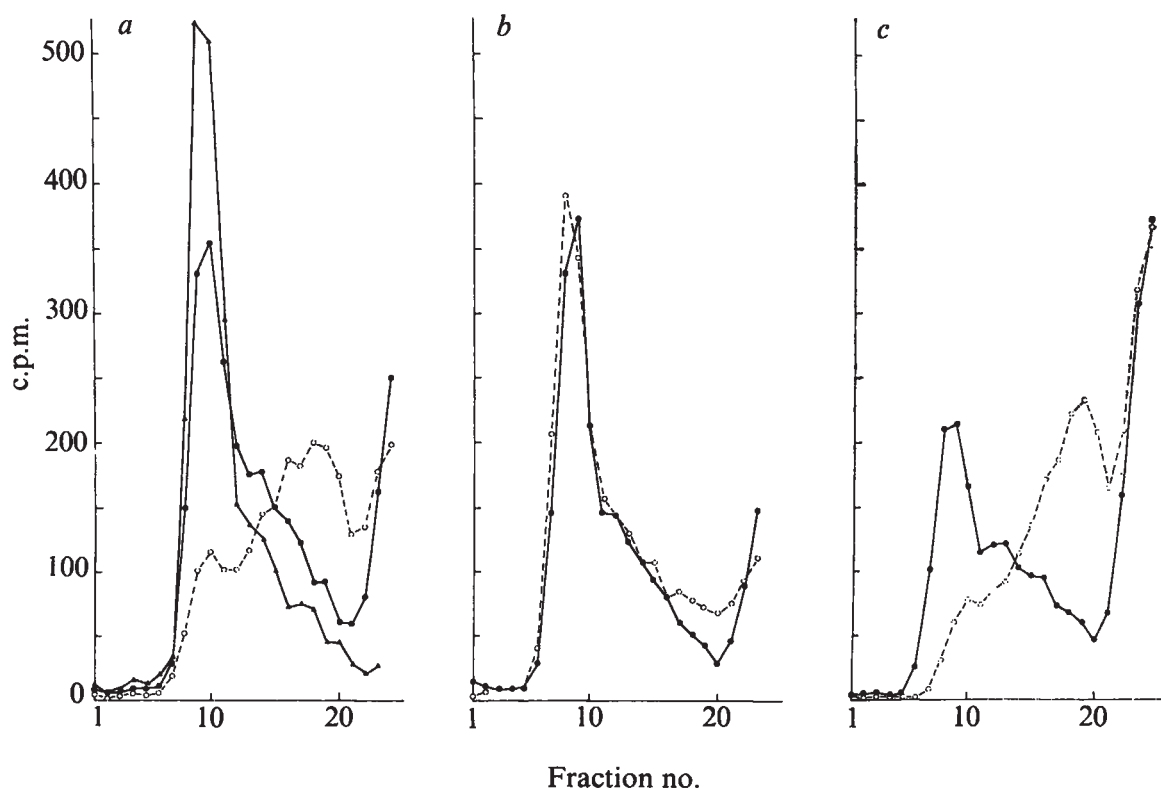


Fig. 1 The preparation of $S30_C$, $S30_{INT}$, reovirus dsRNA, 3H -labelled reovirus mRNA (enriched in medium size class mRNAs) and the procedures for the reovirus mRNA degradation assay have been described before⁹. The reaction mixtures included $30 A_{260} ml^{-1}$ of $S30$ which had been passed through Sephadex G-25, $5 \mu g ml^{-1}$ of reovirus dsRNA, $50 \mu g ml^{-1}$ of 3H -labelled reovirus mRNA as well as all the components of an *in vitro* protein synthesising system⁹, with the following omissions: *a*, amino acids; *b*, amino acids and all the energy source (namely, ATP, GTP, CTP, phosphoenolpyruvate, creatine phosphate, creatine kinase); *c*, amino acids and all the energy source except ATP. The reaction mixtures were incubated at $30^\circ C$ for 30 min. The reactions were terminated by the addition of buffer A (0.1 M NaCl, 10 mM Tris-Cl (pH 7.5), 1 mM EDTA) containing 0.5% SDS. Proteins were removed by extraction with equal volumes of phenol. The aqueous layers (containing RNA) were layered on 12.5-ml linear sucrose gradients (7–25% w/v) in buffer A. The gradients were centrifuged at $2^\circ C$ at 39,000 r.p.m. for 12 h in the SB283 rotor in an IEC B60 ultracentrifuge. Fractions were collected and the radioactivity in each was determined. Δ , Unincubated reaction mixture with $S30_{INT}$; $\bullet$, reaction mixture with $S30_C$; $\circ$, reaction mixture with $S30_{INT}$.

phorylation by ATP of at least two proteins in $S30_{INT}$. Here we describe further characteristics of the faster degradation of reovirus mRNA in $S30_{INT}$ supplemented with dsRNA. ATP is required, in addition to dsRNA, for the acceleration of mRNA degradation in $S30_{INT}$ (ref. 11), and we show that this phenomenon can be divided into two phases. In the first phase (activation) both dsRNA and ATP, but not reovirus mRNA, are required; in the second phase (endonuclease action), in which added reovirus mRNA is degraded, it seems that neither dsRNA nor ATP has to be present.

Previously we have observed the faster degradation of reovirus mRNA in the presence of dsRNA in reaction mixtures containing $S30_{INT}$ as well as all components used for *in vitro* protein synthesis^{9,10}. As Fig. 1*a* shows, faster degradation is also manifested in a reaction mixture not supplemented with amino acids. It is not manifested when neither amino acids nor the energy source is present in the reaction mixture (Fig. 1*b*), but it is manifested when ATP is added (Fig. 1*c*). The commercially available ATP can be replaced by a further purified ATP preparation containing less than 0.1% ADP as the only impurity detected by chromatography and ultraviolet absorption.

The existence of two phases in the phenomenon of faster reovirus mRNA degradation in $S30_{INT}$ is revealed in the experiments shown in Figs 2 and 3. The addition to the reaction mixture of the various components, namely dsRNA, ATP, RNase III (an enzyme hydrolysing dsRNA, but not single-stranded RNA)^{12,13}, glucose and hexokinase (transforming ATP to ADP), and reovirus mRNA, was staggered in these experiments. The results in Fig. 2*a* and *b*

verify that dsRNA is required, in addition to ATP, for the faster mRNA degradation. The faster mRNA degradation is also manifested when both ATP and dsRNA are added at 0 min but dsRNA is degraded after 5 min by the addition of RNase III (Fig. 2*c*). To prove that an RNase III-treated dsRNA cannot activate $S30_{INT}$ for faster mRNA degradation it is shown that there is no enhanced mRNA degradation: (1) if dsRNA and RNase III are added simultaneously at 0 min and ATP as well as reovirus mRNA are added later (Fig. 2*d*), or (2) if dsRNA, RNase III and ATP are added together at 0 min and reovirus mRNA is added later (data not shown). These results do not rule out the possibility that segments of dsRNA might attach to protein during the 5-min incubation (in the absence of RNase III) and thus be protected from hydrolysis by RNase III added later.

The results in Fig. 3*a* and *b* verify that ATP is required, in addition to dsRNA for the faster reovirus mRNA degradation. The activated state persists even after ATP is removed by the addition of hexokinase and glucose (Fig. 3*c*). Hexokinase and glucose are effective in removing ATP since no activated state is reached if ATP, hexokinase and glucose are added simultaneously at 0 min and dsRNA and reovirus are added later (Fig. 3*d*). We have also established that a reaction mixture, which has been incubated with dsRNA and ATP (from 0 min), retains its activated state even if dsRNA is hydrolysed (by RNase III added at 5 min) and ATP is removed (by hexokinase and glucose added at 5 min) (data not shown).

These results indicate that dsRNA and ATP are needed in the activation phase but seem not to be needed in the

endonuclease action phase. The role (if any) of the dsRNA and ATP-promoted protein phosphorylation in S30_{INT} (ref. 10) in the process of activation is under investigation. It remains to be seen whether the faster degradation of reovirus mRNA in S30_{INT} is a result of increased endonuclease activity or of increased susceptibility of the reovirus mRNA to endonuclease action. An increased susceptibility of the mRNA may arise if dsRNA diminishes ribosome binding to mRNA and polysome formation more in S30_{INT} than in S30_C (ref. 6). The hypothesis of increased mRNA susceptibility as the basis of the phenomenon is made unlikely by the finding that the protein synthesis inhibitors (sparsomycin and edeine) do not diminish the difference between S30_{INT} and S30_C in the rate of reovirus mRNA degradation in the presence of dsRNA (to be published). The addition of dsRNA and ATP to S30_{INT} affects the rate of degradation of some RNAs much more than that of others. It accelerates the degradation of reovirus mRNAs from the large size class much more than that of those from the medium or small size classes. Furthermore, it increases the rate of degradation of an unfractionated mRNA preparation from Ehrlich ascites tumour cells much more than that of either Ehrlich ascites tumour cell ribosomal RNA or mouse globin mRNA (to be published). It may be relevant to note that

Fig. 2 The reaction mixtures included all the components as described for Fig. 1b except that 33 $\mu\text{g ml}^{-1}$ of reovirus mRNA was added only after 35 min incubation and all reactions were performed with S30_{INT} only. ATP (1 mM), reovirus dsRNA (2 $\mu\text{g ml}^{-1}$) and 25 U ml^{-1} of RNase III (a gift from Dr R. Crouch) were added as indicated at the times specified. The reaction mixtures were incubated at 30 °C for a total of 50 min. *a*, -dsRNA + ATP (0 min); *b*, + dsRNA (0 min) + ATP (0 min); *c*, + dsRNA (0 min) + ATP (0 min) + RNase III (5 min); *d*, + dsRNA (0 min) + RNase III (0 min) + ATP (30 min).

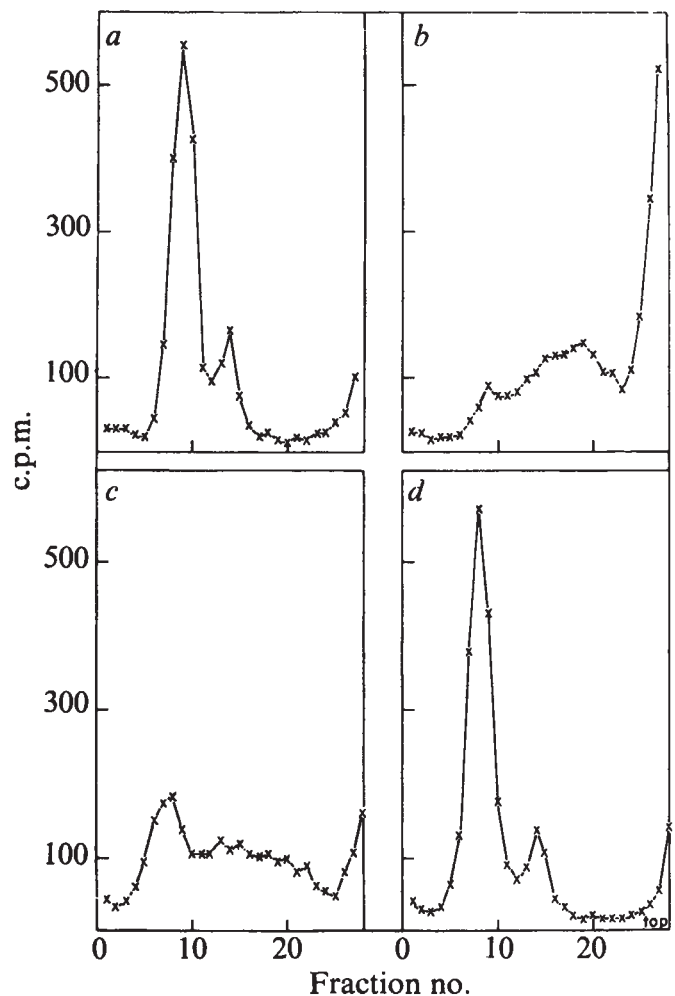
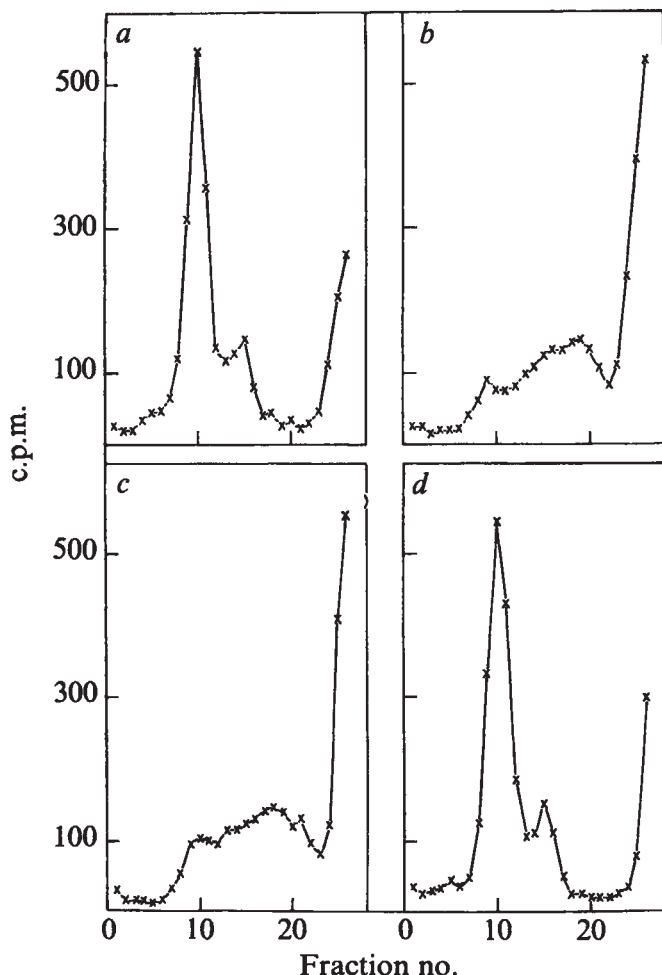


Fig. 3 The conditions were similar to Fig. 2 except that 10 mM glucose and 10 U ml^{-1} of hexokinase (Sigma) were added in *c* and *d* at 5 and 0 min, respectively. It was shown in control experiments that 10 mM glucose, or 10 U ml^{-1} of hexokinase, or 1 mM glucose-6-phosphate added individually at 0 min (together with dsRNA and ATP) did not impair the activation of the reaction mixture for faster mRNA degradation (data not shown). *a*, + dsRNA (0 min) - ATP; *b*, + dsRNA (0 min) + ATP (0 min); *c*, + dsRNA (0 min) + ATP (0 min) + glucose and hexokinase (5 min); *d*, + ATP (0 min) + glucose and hexokinase (0 min) + dsRNA (30 min).



viral RNA accumulation has been reported to be impaired in interferon-treated cells infected with different viruses¹⁴⁻¹⁷.

We thank R. Stout for analysing our ATP sample. This study was supported by grants from the NIH.

G. C. SEN
B. LEBLEU
G. E. BROWN
M. KAWAKITA
E. SLATTERY
P. LENGUEL

Department of Molecular Biophysics and
Biochemistry,
Yale University,
New Haven, Connecticut 06520

Received July 22; accepted September 27, 1976.

- 1 Finter, N. B. (ed.), *Interferons and interferon inducers*, 2 (North-Holland, Amsterdam, 1973).
- 2 Gupta, S. L., Sopori, M. L., and Lengyel, P., *Biochem. biophys. Res. Commun.*, **57**, 763-770 (1974).
- 3 Sen, G. C., et al., *Biochem. biophys. Res. Commun.*, **65**, 427-434 (1975).
- 4 Sen, G. C., et al., *J. Virol.*, **17**, 191-203 (1976).
- 5 Friedman, R. M., et al., *J. Virol.*, **10**, 1184-1198 (1972).
- 6 Kerr, I. M., Brown, R. E., and Ball, L. A., *Nature*, **250**, 57-59 (1974).
- 7 Content, J., Lebleu, B., Zilberstein, A., Berissi, H., and Revel, M., *FEBS Lett.*, **41**, 125-130 (1974).

- ⁸ Samuel, C. E., and Joklik, W. K., *Virology*, **58**, 476-491 (1974).
⁹ Brown, G. E., et al., *Biochem. biophys. Res. Commun.*, **69**, 114-122 (1976).
¹⁰ Lebleu, B., Sen, G. C., Shaila, S., Cabrer, B., and Lengyel, P., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 3107-3111 (1976).
¹¹ Brown, G. E., et al., *Fedn Proc.*, **35**, 1414 (1976).
¹² Robertson, H. D., Webster, R. E., and Zinder, N. D., *J. biol. Chem.*, **243**, 82-91 (1968).
¹³ Crouch, R. J., *J. biol. Chem.*, **249**, 1314-1316 (1974).
¹⁴ Oxman, M. H., and Levin, M. J., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 299-302 (1971).
¹⁵ Marcus, P. I., Engelhardt, D. L., Hunt, J. M., and Seckellick, M. J., *Science*, **174**, 593-598 (1971).
¹⁶ Manders, E. K., Tilles, J. G., and Huang, A. S., *Virology*, **49**, 573-581 (1972).
¹⁷ Bialy, H. S., and Colby, C., *J. Virol.*, **9**, 286-289 (1972).

Structure of porcine pancreatic prephospholipase A₂

PHOSPHOLIPASE A₂ (EC 3.1.1.4) catalyses the hydrolysis of the fatty acid ester bond at the 2-position of 1,2 diacyl sn-phosphoglycerides, and has been purified from many sources, such as mammalian pancreas^{1,2}, snake venom³ and bee venom⁴. In mammals, the enzyme is produced as a precursor, which trypsin can convert into active phospholipase A₂ by splitting the peptide bond Arg 7-Ala 8, with the removal of the N-terminal heptapeptide from the molecule⁵. The activity of the different phospholipases A₂ depends on

the physical state of the substrate. The enzyme from snake venom is highly active when the substrate is tightly packed in a membrane. Mammalian phospholipases cannot attack membranes; they can degrade the lecithin molecules most effectively when these are more loosely packed in micelles. All phospholipases also have some activity towards substrates in a monomeric solution. The precursors from mammalian pancreas phospholipases are also active towards the monomeric substrate molecules, but inactive to substrates in micellar form⁶. The primary structure of many of the phospholipases A₂ is known. They have a single polypeptide chain of approximately 130 amino acid residues and six or seven disulphide bridges⁷⁻¹⁰. There is a high degree of homology between all the enzymes. We report here the three-dimensional structure of one of these enzymes, prephospholipase A₂ from porcine pancreas.

Crystallisation was accomplished as follows. The enzyme was dissolved in a 0.05 M Tris-maleate buffer, pH 7.2, with 5 mM CaCl₂. To 50 µl of this solution was added 10-15 µl of methanol. After 1 week crystals had grown to sizes up to 0.5×0.5×0.5 mm. The space group of the crystal form was P3₁21 with *a*=*b*=69.8 Å and *c*=67.6 Å. There was one molecule in the asymmetric unit and the solvent content

	8	10	12	14	16	18	20	22	24	26	28	30	32	34	
a.	H ₂ N Ala	Leu	Trp	Gln Phe	Arg Ser Met	Ile	Lys Cys Ala	Ile Pro	Gly Ser His	Pro Leu Met	Asp Phe	Asn Asn	Tyr Gly Cys		
b.	H ₂ N Asp	Leu	Thr	Gln Phe	Gly Asn Met	Ile	Asn ---	--- Lys	Met Gly Glu	Ser Val Phe	Asp Tyr	Ile Tyr	Tyr Gly Cys		
c.	H ₂ N Asn	Leu	Tyr	Gln Phe	Lys Asn Met	Ile	His Cys Thr	Val Pro	Asn Arg ---	Pro Trp Trp	His Phe	Ala Asn	Tyr Gly Cys		
d.	H ₂ N Asn	Leu	Tyr	Gln Phe	Lys Asn Met	Ile	Gln Cys Thr	Val Pro	Asn Arg ---	Ser Trp Trp	His Phe	Ala Asn	Tyr Gly Cys		
e.	H ₂ N Asn	Leu	Tyr	Gln Phe	Lys Asn Met	Ile	His Cys Thr	Val Pro	Asn Arg ---	Ser Trp Trp	His Phe	Ala Asn	Tyr Gly Cys		
f.	H ₂ N Asn	Leu	Tyr	Gln Phe	Lys Asn Met	Ile	Lys Cys Thr	Val Pro	Ser Arg ---	Ser Trp Trp	His Phe	Ala Asn	Tyr Gly Cys		
g.	H ₂ N Asn	Leu	Val	Gln Phe	Ser Tyr Leu	Ile	Gln Cys Ala	Asn His	Gly Lys Arg	Pro Thr Trp	His Tyr	Met Asp	Tyr Gly Cys		
	36	38	40	42	44	46	48	50	52	54	56	58	60		
	Tyr Cys Gly	Leu	Gly Gly	Ser Gly	Thr Pro	Val Asn	Glu Leu	Asn Arg	Cys Cys	Glu Thr	His Asp	Asn Cys	--- Tyr	Arg Asp	
	Tyr Cys Gly	Trp	Gly Gly	Lys Gly	Lys Pro	Ile Asp	Ala Thr	Asp Arg	Cys Cys	Phe Val	His Asp	--- Cys	Cys Tyr	Gly Lys	
	Tyr Cys Gly	Arg	Gly Gly	Lys Gly	Thr Pro	Val Asp	Asp Leu	Asp Arg	Cys Cys	Gln Ile	His Asp	Lys Cys	--- Tyr	Asp Glu	
	Tyr Cys Gly	Arg	Gly Gly	Ser Gly	Thr Pro	Val Asp	Asp Leu	Asp Arg	Cys Cys	Gln Ile	His Asp	Asn Cys	--- Tyr	Gly Glu	
	Tyr Cys Gly	Arg	Gly Gly	Ser Gly	Thr Pro	Val Asp	Asp Leu	Asp Arg	Cys Cys	Gln Thr	His Asp	Asn Cys	--- Tyr	Ser Asp	
	Tyr Cys Gly	Ala	Gly Gly	Ser Gly	Thr Pro	Val Asp	Glu Leu	Asp Arg	Cys Cys	Lys Ile	His Asp	Asp Cys	--- Tyr	Asp Glu	
	62	64	66	68	70	72	74	76	78	80	82	84	86	88	
	Ala Lys Asn	Leu Asn	Asp Ser	Cys Lys	Phe Leu	Val Asp	Asn Pro	Tyr Thr	Glu Ser	Tyr Ser	--- Tyr	Cys Ser	Ser Ser	Asn Thr	
	Met Gly	--- Thr	Tyr Asp	Thr ---	Lys Trp	Thr Ser	Tyr Asn	--- Tyr	--- Glu	Ile ---	--- ---	--- Gln	--- Asn	Gly	
	Ala Glu	--- Lys	Ile Ser	Gly Cys	Trp Pro	Tyr Ile	Lys Thr	--- Tyr	Thr ---	--- Tyr	Glu ---	Ser Cys	Gln Gly	--- Thr	
	Ala Glu	--- Lys	Ile Ser	Gly Cys	Trp Pro	Tyr Ile	Cys Thr	--- Tyr	Thr ---	--- Tyr	Glu ---	Ser Cys	Gln Gly	--- Thr	
	Ala Glu	--- Lys	Ile Ser	Gly Cys	Trp Pro	Tyr Ile	Lys Thr	--- Tyr	Ser ---	--- Tyr	Asp ---	Ser Cys	Gln Gly	--- Thr	
	Ala Glu	--- Lys	Ile Ser	Gly Cys	Arg Pro	Tyr Phe	Lys Thr	--- Tyr	Ser ---	--- Tyr	Asp ---	Cys Thr	Lys Gly	Lys	
	Ala Lys	Ala Gly	Lys Lys	Gly Cys	--- ---	--- ---	--- Phe	Pro Lys	Met Ser	Ala Tyr	Asp Tyr	Tyr Tyr	Cys Gly	Glu Asn ---	
	90	92	94	96	98	100	102	104	106	108	110	112	114	116	
	Glu Ile Thr	Cys	Asn Ser	Lys Asn	Asn Ala	Cys Glu	Ala Phe	Ile Cys	Asn Cys	Asp Arg	Asn Ala	Ala Ala	Ile Cys	Phe Ser	Lys
	Gly Ile Asp	Cys	Asp Glu	Asp Pro	Gln Lys	--- Lys	Glu Leu	--- Cys	Glu Cys	Asp Arg	Val Ala	Ala Ala	Ile Cys	Phe Phe	Ala Asn
	Leu Thr	--- Cys	Lys Asp	Gly Gly	Lys ---	Cys Ala	Ala Ser	Val Cys	Asp Cys	Asp Arg	Val Ala	Ala Ala	Asn Cys	Phe Phe	Ala Arg
	Leu Thr Ser	Cys	Gly Ala	Asn Asn	Lys ---	Cys Ala	Ala Ser	Val Cys	Asp Cys	Asp Arg	Val Ala	Ala Ala	Asn Cys	Phe Phe	Ala Arg
	Leu Thr Ser	Cys	Gly Ala	Asn Asn	--- Cys	Ala Ala	Ser Val	Cys Cys	Asp Cys	Asp Arg	Val Ala	Ala Ala	Asn Cys	Phe Phe	Ala Arg
	Leu Thr	--- Cys	Lys Glu	Gly Asn	Asn Glu	Cys Ala	Ala Phe	Val Cys	Lys Cys	Asp Arg	Leu Ala	Ala Ala	Ile Cys	Phe Phe	Ala Gly
	Gly Pro Tyr	Cys	Arg Asn	Ile Lys	Lys Lys	Cys Leu	Arg Phe	Val Cys	Asp Cys	Asp Val	Glu Ala	Ala Ala	Phe Cys	Phe Phe	Ala Lys
	118	120	122	124	126	128	130	132							
	Ala Pro Tyr	Asn Lys	Glu His	Lys Asn	--- Leu	Asn Thr	Lys Lys	Tyr ---	Cys OH						
	Asn Arg	Asn Thr	Tyr Asn	Ser Asn	Tyr ---	Phe Gly	His Ser	Ser Ser	Lys Cys	Thr Gly	Thr Glu	Gln Cys	OH		
	Ala Thr Tyr	Asn Asp	Lys Asn	--- Tyr	Asn Ile	Asp Phe	Asn Ala	Arg ---	Cys Gln	OH					
	Ala Thr Tyr	Asn Asp	Lys Asn	--- Tyr	Asn Ile	Asp Phe	Asn Ala	Arg ---	Cys Gln	OH					
	Ala Pro Tyr	Ile Asp	Lys Asn	--- Tyr	Asn Ile	Asp Phe	Asn Ala	Arg ---	Cys Gln	OH					
	Ala His Tyr	Asn Asp	Asn Asn	Tyr ---	Ile Asp	Leu ---	Ala Arg	His Cys	Gln OH						
	Ala Pro Tyr	Asn Asn	Ala Asn	Trp Asn	--- Ile	Asp Thr	Lys Lys	Arg ---	Cys Gln	OH					

Fig. 1 Amino acid sequence of proteins with phospholipase A₂ activity. The numbering is from the porcine prephospholipase A₂ sequence. a, Porcine phospholipase A₂. Some modifications in the sequence, as given in ref. 7, will be published shortly by De Haas et al. b, *Bitis gabonica* phospholipase A₂ (8); c, d and e, *Naja melanoleuca* phospholipase A₂, respectively, fractions DE-I, DE-II and DE-III (9); f, *Hemachatus haemachatus* phospholipase A₂, fraction DE-I (9); g, notexin (10).

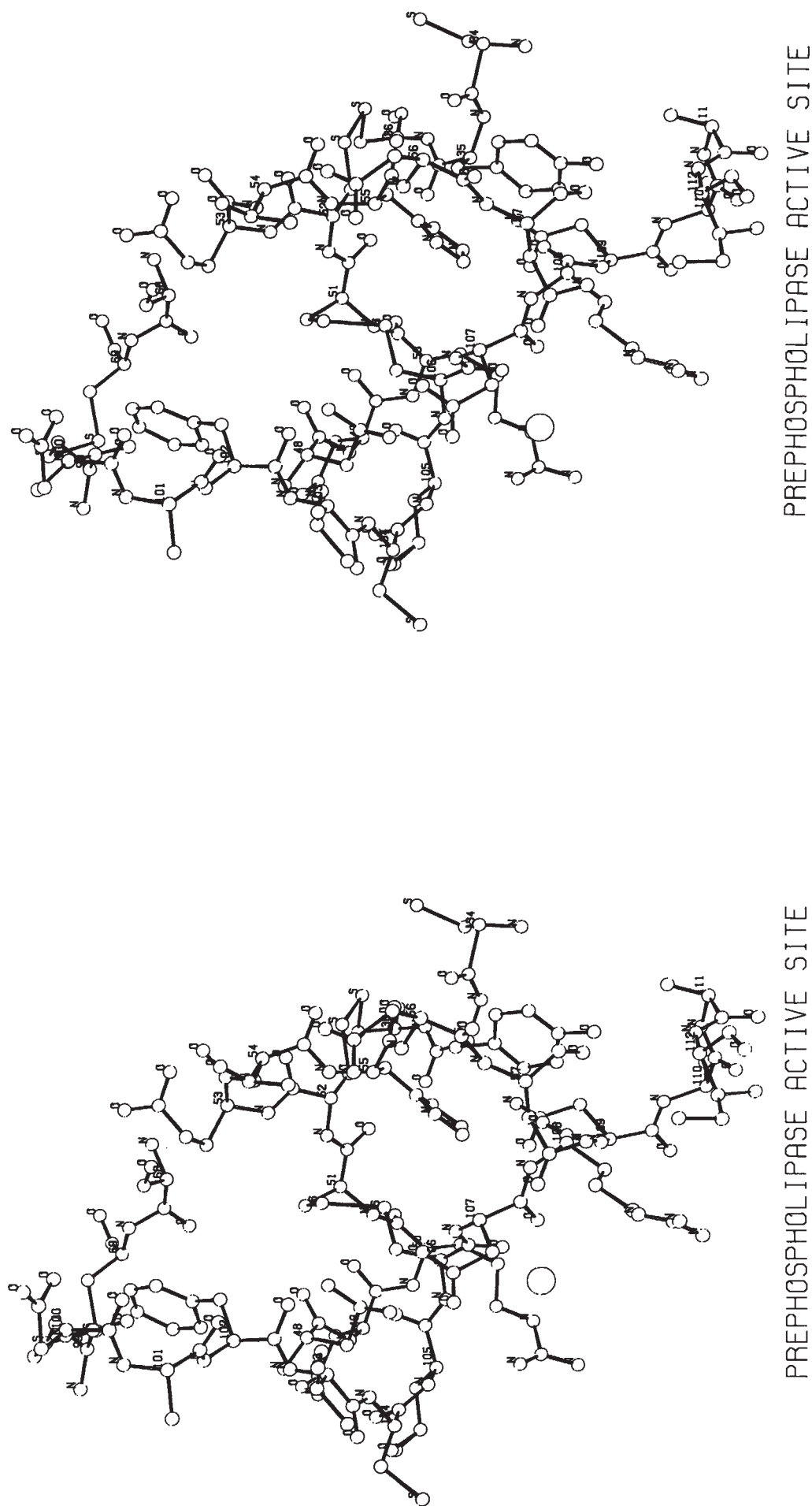


Fig. 2 Stereo plot of the main chain and disulphide bridges of porcine pancreatic prephospholipase A₂.

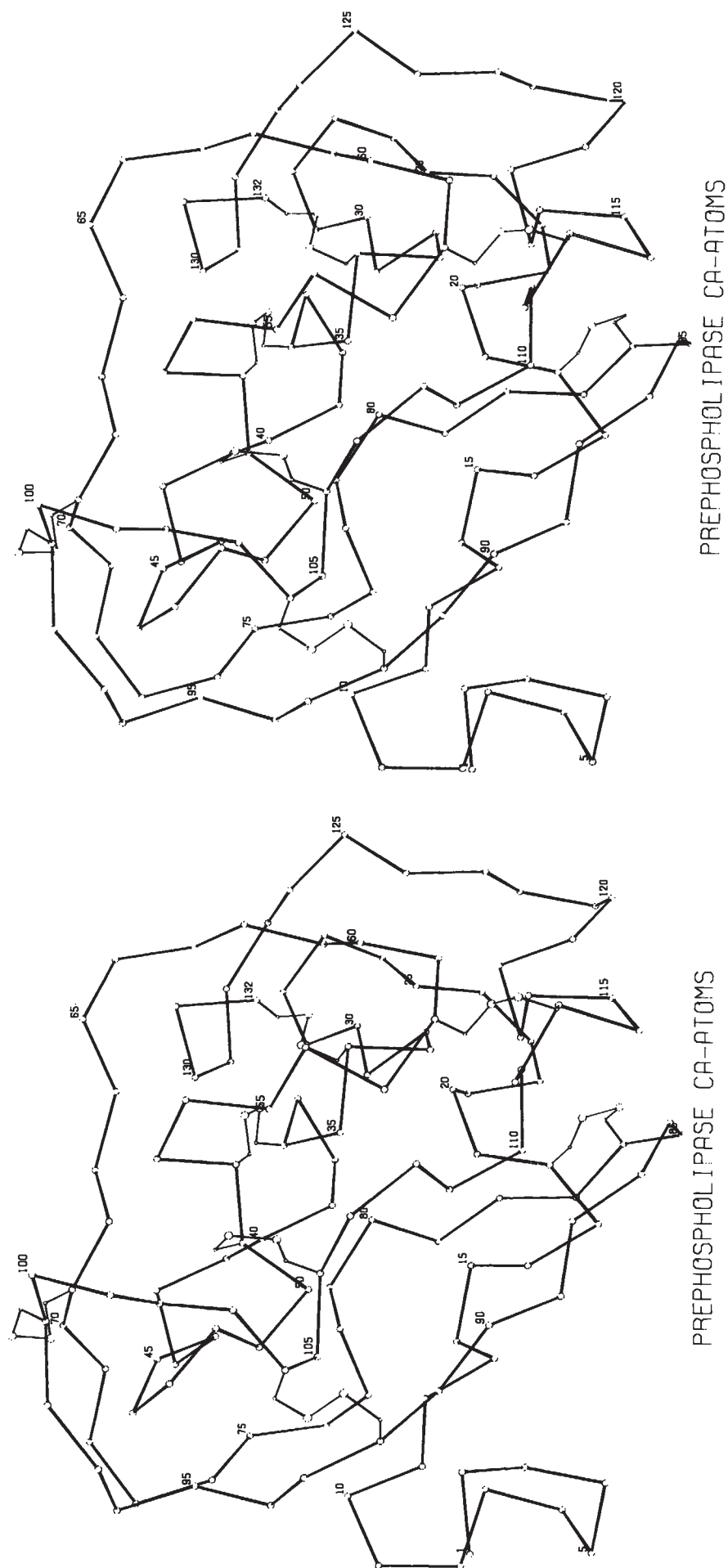


Fig. 3 Stereo plot of the active site region of porcine pancreatic prephospholipase A₂.

of the crystals was 62%. For structure determination we used one heavy atom derivative— $K_3UO_2F_6$ —with the anomalous scattering data of the uranyl atom (the SIRAS method). An electron density map was calculated to a resolution of 3 Å. This map—though of mediocre quality—could be interpreted easily with the aid of the known amino acid sequence⁷.

The polypeptide chain in the prephospholipase A_2 molecule consists of 132 amino acid residues, 14 of them involved in disulphide bridges. These seven bridges provide an internal check on the correctness of the interpretation of the electron density map, because the polypeptide chain always has to return to a bridge at a known position in the amino acid sequence.

The molecule has the shape of a rectangular box, with dimensions of approximately $25 \times 28 \times 35$ Å (Fig. 2). It contains neither regular α helices nor β -pleated sheets; only two parts are in a distorted helix conformation—residues 11–17 and 36–41. Four disulphide bridges are more or less in a plane through the middle of the molecule and three are at the surface. The N-terminal heptapeptide in prephospholipase has the sequence Glu–Glu–Gly–Ile–Ser–Ser–Arg. It shows up as a bent tail at the molecular surface. During the activation of prephospholipase by the action of trypsin the peptide bond Arg 7–Ala 8 is split. This bond is in an exposed position on the surface of the molecule.

His 55 is part of the active site of porcine (pre-)phospholipase A_2 (ref. 11). Incorporation studies, using ^{14}C -labelled *p*-bromophenacyl bromide showed that loss of enzyme activity was accompanied by the incorporation of 1.1 ^{14}C -*p*-bromophenacyl residues and by the loss of 1 histidyl residue per mol of protein. Divalent metal ions (Ca and Ba), however, protect the enzyme against this inactivation. In the electron density map, we find in the neighbourhood of His 55 a peak of density that can be interpreted as a Ca ion. It is close to the side chains of Asp 107 and Asn 57. Ca ions are also an absolute requirement for crystallisation. The crystals crack when the Ca ions are removed by addition of EDTA. When calcium binds to the protein, the spectrum in the ultraviolet region changes¹². The difference spectrum shows a perturbation of one or more tyrosine and histidine residues. One tyrosine residue, Tyr 35, is close to His 55 and the Ca-binding site (Fig. 3).

The molecular model suggests a binding mode for monomers of the substrate. It is reasonable to assume that the phosphate moiety of the substrate is bound to Arg 108 and the Ca ion. This would be similar to staphylococcal nuclease where the 5'-phosphate group of the substrate is bound specifically to the enzyme by two guanidino groups with the Ca ion 4 Å from the nearest phosphate oxygen atom¹³. In phospholipase A_2 Arg 108 is always present, except for notexin¹⁰, but the catalytic action for the latter enzyme seems to be different. The carbonyl oxygen of the ester bond to be split could well point towards the imidazole ring of His 55. Then the hydrophobic part of the lecithin molecule could fold along a highly hydrophobic surface part of the enzyme. This surface part consists of the side chains Leu 48, Cys 51–Cys 106, Thr 54, Ile 103, Phe 102 and Cys 69–Cys 99.

To predict a mechanism for the esterase activity of phospholipase we made a comparison with other ester splitting enzymes. It is well known that proteolytic enzymes have esterase as well as proteolytic activity and that the mechanism for splitting an ester bond is essentially the same as for splitting a peptide bond. X-ray crystallographic analysis of several proteolytic enzymes has provided a picture of the catalytic process in the active site.

The catalytic action of the serine enzymes starts with nucleophilic attack of the serine OH on the carbonyl carbon of the substrate^{14,15}. A tetrahedral intermediate, which has a negative charge on its carbonyl oxygen, is formed. The

Table 1 Functional groups in esterases

Enzyme	Nucleophile	Proton donor	Stabiliser
Serine enzymes	Ser OH	Imidazole H ⁺	Amide NHs
Papain	Cys S ⁻	Imidazole H ⁺	Amide NHs
Carboxypeptidase	COO ⁻ (+H ₂ O)	Tyr OH	Zinc
Thermolysin	COO ⁻ (+H ₂ O)	Imidazole H ⁺	Zinc
Phospholipase A_2	COO ⁻ (+H ₂ O)	Tyr OH	Imidazole H ⁺

tetrahedral intermediate is stabilised by amide NH groups in the active site wall. The splitting of the peptide or ester bond is accompanied by a protonation of the leaving group at the peptide NH or the ester oxygen atom. The proton donating group in the serine enzymes is the imidazole ring next to the serine, transferring a proton from the serine to the leaving group. As a result of this process the enzyme is acylated. Deacylation proceeds through similar steps, water replacing the leaving group. Thus three functions have a key role in the catalytic mechanism of the serine enzymes (Table 1): (1) the serine OH as a nucleophile; (2) the amide NH groups as stabilisers for the tetrahedral intermediate; and (3) proton donating imidazole.

From an X-ray study of papain derivatives a reasonably reliable picture has been obtained for the catalytic mechanism of this sulphhydryl protease¹⁶. It turned out to be very similar to the action of the serine enzymes. The sulphhydryl group is the nucleophile. The side chain NH₂ of Gln 19 and the main chain NH of Cys 25 contribute to the stabilising function. The imidazole ring of His 159 gives up a proton to the leaving group.

The proteolytic zinc-dependent enzymes, carboxypeptidase A¹⁷ and thermolysin¹⁸, both use the zinc ion as the stabiliser for the carbonyl oxygen. They also have the same nucleophile, which is either the carboxyl group of a Glu residue or a water molecule with its nucleophilicity increased by these carboxyl groups. The proton donating group is different for these two enzymes, being Tyr in carboxypeptidase A and His in thermolysin.

A similar set of three important functional groups should be present in other enzymes with a proteolytic or ester splitting activity. In this way we could assign functions to some side chains in the active site of phospholipase A_2 . Searching for the three functions mentioned above we find Asp 56, His 55 and Tyr 35. These residues are conserved in all other phospholipases sequenced. The position of the three residues in the molecular model suggests that the Asp is the nucleophile, the imidazole ring the stabiliser and that Tyr has a proton donating function. These predictions for phospholipase A_2 should be verified, for example, by crystallographic binding studies and specific chemical modifications.

All phospholipases A_2 with known amino acid sequence show a striking homology, especially for those parts of the polypeptide chain which are in the active site region. His 55, Tyr 35, Asp 56 and Asp 107 are absolutely conserved: the conservative part 32–37 has two cysteines in S–S bridges with the main chain atoms of Tyr 35 fixed in between. The hydrophobic region shows a functional homology: only Cys 51–Cys 106 is the same in all enzymes; for the other parts a high degree of hydrophobicity is maintained, but not necessarily by identical residues. When, for example, Phe 102 is replaced by a serine residue, threonine at position 54 is always replaced by isoleucine. The most variable parts of the sequence are located on the surface of the molecule, away from the active site region.

The high activity of active phospholipase A_2 towards the micellar substrate cannot yet be explained.

We thank Professor De Haas for the protein material and Mrs G. Drent for technical assistance. We thank Drs M.

Pierrot and H. Reynaers for crystallising the enzyme. This work was supported in part by the Netherlands Foundation for Chemical Research with financial aid from the Netherlands Organisation for the Advancement of Pure Research.

J. DRENTH
C. M. ENZIG
K. H. KALK
J. C. A. VESSIES

Laboratory for Structural Chemistry,
Zernikelaan, Paddepoel,
Groningen, The Netherlands

Received July 19; accepted September 21, 1976.

- ¹ De Haas, G. H., Postema, N. M., Nieuwenhuizen, W., and Van Deenen, L. L. M., *Biochim. biophys. Acta*, **159**, 103–117 (1968).
- ² Dutih, C. I., Van Doren, P. J., Verheul, F. E. A. M., and De Haas, G. H., *Eur. J. Biochem.*, **53**, 91–97 (1975).
- ³ Wells, M. A., *Biochemistry*, **10**, 4074–4078 (1971).
- ⁴ Shipolini, R. A., et al., *Eur. J. Biochem.*, **20**, 459–468 (1971).
- ⁵ De Haas, G. H., Postema, N. M., Nieuwenhuizen, W., and Van Deenen, L. L. M., *Biochim. biophys. Acta*, **159**, 118–129 (1968).
- ⁶ Pieterse, W. A., Vidal, J. C., Volwerk, J. J., and De Haas, G. H., *Biochemistry*, **13**, 1455–1460 (1974).
- ⁷ De Haas, G. H., et al., *Biochim. biophys. Acta*, **221**, 31–53, 54–61 (1970).
- ⁸ Botes, D. P., and Viljoen, C. C., *J. biol. Chem.*, **249**, 3827–3835 (1974).
- ⁹ Joubert, J. J., *Eur. J. Biochem.*, **52**, 539–554 (1975).
- ¹⁰ Halpert, J., and Eaker, D., *J. biol. Chem.*, **250**, 6990–6997 (1975).
- ¹¹ Volwerk, J. J., Pieterse, W. A., and De Haas, G. H., *Biochemistry*, **13**, 1446–1454 (1974).
- ¹² Pieterse, W. A., Volwerk, J. J., and De Haas, G. H., *Biochemistry*, **13**, 1439–1445 (1974).
- ¹³ Cotton, F. A., Bier, C. J., Day, V. W., Hazen, E. E., and Larsen, S., *Cold Spring Harb. Symp. quant. Biol.*, **36**, 243–249 (1971).
- ¹⁴ Robertus, J. D., Kraut, J., Alden, R. A., and Birktoft, J. J., *Biochemistry*, **8**, 4293–4303 (1972).
- ¹⁵ Henderson, R., Wright, C. S., Hess, G. P., and Blow, D. M., *Cold Spring Harb. Symp. quant. Biol.*, **36**, 63–70 (1971).
- ¹⁶ Drenth, J., Swen, H. M., and Kalk, K. H., *Biochemistry*, **15**, 3731–3738 (1976).
- ¹⁷ Hartsuck, J. A., and Lipscomb, W. N., in *The Enzymes*, **3**, third ed. (edit. by Boyer, P. D.), 1–56 (Academic, New York and London, 1971).
- ¹⁸ Matthews, B. W., et al., *Nature new Biol.*, **238**, 41–43 (1972).

Organisation of the polypeptide chains in mammalian keratin

THE keratins are insoluble intracellular α -fibrous proteins which represent the major structural proteins of epidermis. X-ray diffraction studies of keratins in the early 1950s^{1,2} yielded a pattern since shown to be typical of intact epidermal tissue and epidermal α keratins. This pattern, which consists of a meridional arc at 5.15 Å and an equatorial reflection at 9.8 Å, was suggested to be indicative of polypeptide chains packed in a partially α -helical coiled coil configuration. Crick³ suggested that a three-stranded model was appropriate for α keratins, and although two-stranded and seven-stranded models have also been proposed, the triple α -helical coiled coil configuration has been a dominant concept in keratin structure⁴. Rudall made the initial contribution of epidermal keratin chemistry by using urea buffers to extract and characterise the α -fibrous proteins from cow snout epidermis. Matoltsy subsequently showed that an α -fibrous protein, which he designated prekeratin, could also be isolated by extraction with citrate buffer, pH 2.65, and purified by isoelectric precipitation. His data indicated that the prekeratin was an α -helical, rod-like molecule with a molecular weight of 640,000 (ref. 5). Detailed studies^{6–8} have shown that epidermal keratins consist of several different polypeptide chains, and we describe here experiments which suggest how these chains may be organised in the native keratin molecule.

Prekeratin, the epidermal keratin, was prepared according to Matoltsy⁵, and the sodium dodecyl sulphate (SDS)–polyacrylamide electrophoretic pattern obtained using the discontinuous system of Neville⁹ is shown in Fig. 1. The relative proportions of each of the bands were determined by scanning the gel to measure the amounts of Coomassie blue dye bound to each chain. After taking into account that the amount of dye binding is proportional to the molecular weights of the chains, it can be shown that the three major

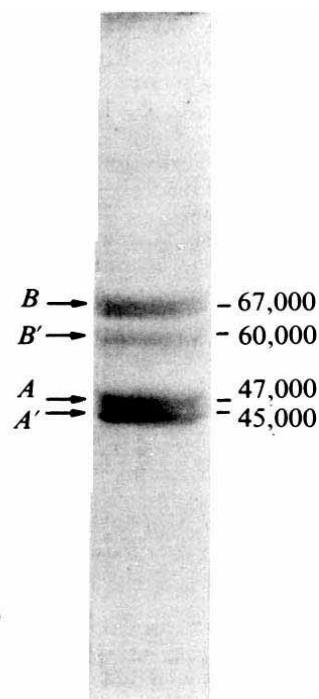


Fig. 1 SDS–polyacrylamide electrophoresis of keratin from cow snout epidermis. The gels are 7% acrylamide 2% bisacrylamide and run at pH 9.5 using the discontinuous gel system described by Neville⁹. Gels were run at 2 m per gel for 2 h at 25 °C and fixed and stained with Coomassie blue in 7.5% acetic acid (v/v) and 5% methanol (v/v). The direction of migration is downward. Samples were equilibrated in 8 M urea, 1% SDS and 1% mercaptoethanol for 1 h at 50 °C and dialysed against 1,000 volumes of upper gel buffer before loading on gels.

bands, designated B, A, A', are present in approximately equal amounts. A fourth component B' is also present but in much smaller amounts than B, A or A'. Having isolated the major polypeptide chains of keratin on DEAE cellulose as previously described¹⁰, we could perform X-ray diffraction studies on fibres pulled from solutions of purified chains and from solutions in which the chains had been recombined in the ratios shown in Table 1. Experimental detail is given

Table 1 X-ray diffraction patterns

Sample	Presence of α pattern
Prekeratin	Yes
A	No
B	No
A + A'	No
B + B'	No
A + B	No
2A + B	Yes
A + A' + B	Yes
2A' + B	Not done

Fibres (between 1 and 2 cm long) from the above solutions were prepared for X-ray diffraction in one of three ways: (1), lyophilised protein was dissolved in 80% formic acid and the gelatinous mixture was taken up and stretched between the tips of forceps; (2) solutions of protein in 0.1 M citric acid were dialysed against 0.1 M NH_4HCO_3 , pH 9.0 and the filaments that formed were picked up and stretched between forceps; (3) solutions of protein dissolved in 8 M urea containing 0.01 M Tris at pH 8.5 were dialysed against water and the filaments that formed were also picked up and stretched between forcep tips. The protein content of solutions of purified chains was measured by the Lowry technique before they were combined in the molar ratios shown. Fibres were mounted with a specimen-to-film distance of 1.54 cm and exposed to $\text{CuK}\alpha$ radiation ($\lambda = 1.54 \text{ Å}$) at 40 kV. Finely powdered KCl was dusted on to selected fibres before exposure and the 2.224 Å spacing was used for calibration purposes. Variations of measurements from replicate fibres was less than 0.1 Å. The presence of an α pattern was indicated by a sharp meridional arc at $5.15 \pm 0.1 \text{ Å}$ and an equatorial reflection at $9.8 \pm 0.1 \text{ Å}$.

in the legend. Purified *A* and *B* chains alone did not produce an X-ray diffraction pattern typical of α -fibrous proteins; however when these chains were combined in a ratio of 2:1 (*A* to *B*) an α X-ray pattern indistinguishable from that of either intact keratin or epidermis was obtained. This is interesting in the light of our previous descriptions of these chains which showed that *A* and *A'* and *B* and *B'* were chemically and immunologically related to one another and formed two distinct families (*A* and *B*) of polypeptide chains associated with keratins¹⁰.

Although these studies supported an unlike-chain (*AA'B* or *AA'B'*) hypothesis for the assembly of the keratin molecule, it seemed possible that a series of keratins composed only of the identical individual peptide chains (*AAA*, *A'A'A'*, *BBB* or *B'B'B'*) might exist. In the unlike-chain arrangement each of the three principal chains (*AA'B*) would be packed in a triple α -helical arrangement with a monomer molecular weight of 159,000. This unit could form dimers (molecular weight 318,000) or tetramers (molecular weight 636,000). The minor component *B'* could also form intact keratin molecules *AA'B'*, with a monomer molecular weight of 153,000 which would be present in smaller amounts and constitute a variant form. Such variant forms have been shown to exist for other structural proteins such as collagen¹¹ and fibrinogen¹².

On the other hand, if identical chains combined with themselves the following species would exist in solution: *AAA* (monomer molecular weight 135,000; dimer, 270,000); *A'A'A'* (monomer, 144,000; dimer, 288,000); *B'B'B'* (monomer, 180,000; dimer, 360,000) and *BBB* (monomer, 201,000; dimer, 402,000). It is possible that in the intact keratin molecule these species are the fundamental units from which the molecule is built.

Intact keratin molecules predicted from the unlike-chain and identical-chain hypotheses would give distinctly different SDS electrophoretic distributions as shown in Fig. 2, providing their interchain associations were not disrupted. Disruption of interchain associations was prevented by chemically crosslinking the polypeptide chains of the keratin molecule with formaldehyde using the techniques of Veiss and Drake¹³ and Nold and Cross¹⁴ for crosslinking collagen. The densitometer tracing of SDS electrophoretic patterns of crosslinked prekeratin (Fig. 3) is clearly similar to that of

Fig. 2 Predicted densitometer tracings of the SDS-polyacrylamide electrophoretic patterns of formaldehyde-crosslinked keratin. Solid line, multichain lymphorexis; broken line, single-chain hypothesis.

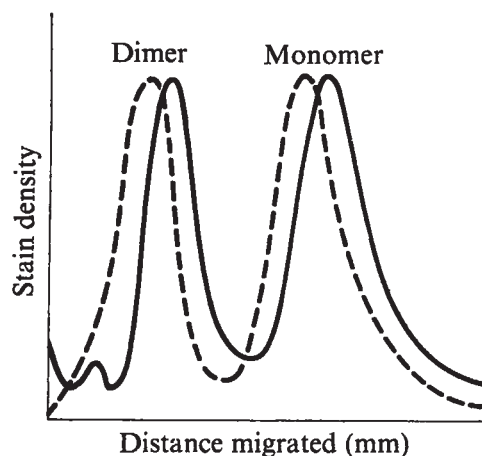
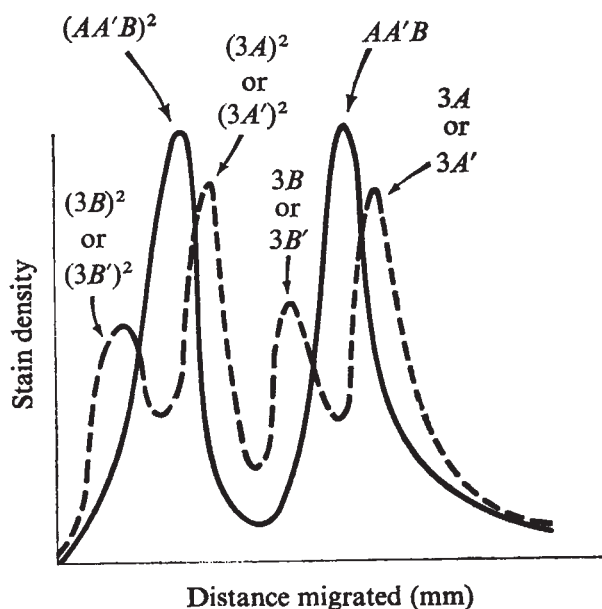


Fig. 3 The solid line represents the densitometer tracing of the SDS-polyacrylamide pattern of formaldehyde-crosslinked prekeratin and the dotted line the theoretical curve for a multichain structure. The amounts of formaldehyde were varied from 1.5 to 6.2% and the times of reaction were varied from 1 h to 3 d to obtain conditions which crosslinked all the available polypeptide chains of keratin but did not yield polymers too heavy to penetrate 5% polyacrylamide gels. The optimal conditions used in all crosslinking experiments reported were as follows. Prekeratin dissolved in 0.1 M acetic acid (1 mg ml⁻¹) was incubated with formaldehyde at a final concentration of 6.2% for 4 h at 4 °C, conditions which were shown not to crosslink solutions of single chain control proteins. After exhaustive dialysis against 0.1 N acetic acid, samples were equilibrated and run as described in Fig. 1, except that 5% acrylamide gels were used. The following molecular weight markers were used: BSA monomer (67,000) and dimer (134,000); phosphorylase A (94,000); λ globulin (150,000).

the predicted unlike-chain molecule comprised of an *A*, *A'* and *B* chain. The variant molecule (comprised of *A*, *A'* and *B'*) is too close in molecular weight to the major species to be seen as a separate peak. The molecular weights obtained for the monomer (144,000) and dimer (280,000) are approximately 10% lower than the theoretical values, which is expected as the crosslinked keratin molecules are not fully accessible to SDS. This may well be analogous to the situation with other proteins such as bovine serum albumin (BSA)¹⁶ where the folded molecule (unreduced) has a lower apparent molecular weight than the unfolded (reduced) species on SDS gels.

Our data indicate that the basic building block of the keratin molecule is composed of two *A* chains and one *B* chain. Skerrow's work¹⁷ in which he isolated a highly helical core, after tryptic digestion of keratin, which consisted of three distinct polypeptide chains, seems to support this concept. It is thought that this basic unit can readily polymerise to form the higher molecular weight units initially reported by Matoltsy and is presumably the fundamental unit of the tonofilaments seen *in vivo*.

We thank Dr Ervin H. Epstein Jr for helpful suggestions.

LORETTA D. LEE
HOWARD P. BADEN

Department of Dermatology,
Harvard Medical School,
Massachusetts General Hospital,
Boston, Massachusetts 02114

Received June 7; accepted September 27, 1976.

- 1 Crick, F. H. C., *Nature*, **170**, 882-883 (1952).
- 2 Pauling, L., and Corey, R. B., *Nature*, **171**, 59-61 (1953).
- 3 Crick, F. H. C., *Acta Crystallogr.*, **6**, 689 (1953).
- 4 Fraser, R. D. B., MacRae, T. P., and Rogers, G. E., *Keratins: Their Composition, Structure and Biosynthesis* (Thomas, Springfield, Illinois, 1972).
- 5 Matoltsy, A. G., in *Biology of Skin and Hair Growth* (edit. by Lyne, A. G., and Short, B. F.), 291-305 (Angus and Robertson, Sydney, 1965).
- 6 Baden, H. P., Goldsmith, L. A., and Fleming, B., *Biochim. biophys. Acta*, **137**, 303-311 (1973).
- 7 Steinert, P. M., *Biochem. J.*, **149**, 39-48 (1975).
- 8 Skerrow, D., *Biochem. biophys. Res. Commun.*, **59**, 1311-1316 (1974).
- 9 Neville, D. M., *J. Biol. Chem.*, **246**, 6328-6334 (1971).
- 10 Lee, L. D., Fleming, B. C., Waitkus, R. F., and Baden, H. P., *Biochim. biophys. Acta*, **412**, 82-90 (1975).

- ¹¹ Epstein, E. N., Jr. and Munderloh, N. H., *J. biol. Chem.*, **250**, 9304–9312 (1975).
¹² Doolittle, R. F., in *Advances in Protein Chemistry*, **27** (edit. by Edsall, J. T., and Richard, F. M.), 1–109 (Academic, New York, 1973).
¹³ Weiss, A., and Drake, M. P., *J. biol. Chem.*, **238**, 2003–2011 (1964).
¹⁴ Noid, J. G., Kang, A. H., and Cross, J., *Science*, **170**, 1096–1098 (1970).
¹⁵ Yphantis, D. A., *Biochemistry*, **7**, 297–317 (1964).
¹⁶ Dunkar, A. K., and Ruekert, R. R., *J. biol. Chem.*, **244**, 5074–5080 (1969).
¹⁷ Skerrow, D., Matoltsy, A. G., Matoltsy, M. N., *J. biol. Chem.*, **248**, 4820–4826 (1973).

'Sandwiched' water molecule between pyrimidine bases and intra-molecular C–H...O hydrogen bonding in 5-nitro-1-(β -D-ribosyluronic acid)-uracil monohydrate

IN the structure of crystalline 5-nitro-1-(β -D-ribosyluronic acid)-uracil monohydrate we have found a novel disposition of a water molecule—'sandwiched' between two parallel pyrimidine rings stacked 6.55 Å apart. The water molecule is held in its place between these rings by a short O–H...OW (2.589 Å) hydrogen bond from the carboxyl group of a uronic acid molecule belonging to an adjacent stack. We have also found that nitro groups adjacent to C–H bonds in the heterocyclic ring of nucleic acid bases apparently polarise these bonds and cause them to take part in strong hydrogen bonds to appropriate donors. The polarisation effect of the nitro group adjacent to the C–H bonds in pyrimidine is perhaps analogous to that of a positive charge adjacent to the C–H bonds in purines¹.

We have been studying the structures of a series of nitro derivatives of nucleic acid bases and nucleosides in connection with a study of radiation damage and sensitisation. To this end, we prepared crystals of 5-nitro-1-(β -D-ribosyluronic acid)-uracil monohydrate by treating uridine with dilute nitric acid², and then slowly evaporating the resulting solution. The crystals so obtained (of $C_9H_9N_3O_9 \cdot H_2O$) are monoclinic, space group $P2_1$, with cell dimensions (at $22 \pm 3^\circ C$): $a = 8.982(1)$, $b = 10.245(1)$, $c = 6.651(1)$ Å, $\beta = 92.27(1)^\circ$, with a density of 1.74 g cm^{-3} (calculated density = 1.743 g cm^{-3}) (this is higher than usually observed for nucleosides). The crystal structure was determined from the intensities of $CuK\alpha$ X-ray 'reflections' measured by a diffractometer out to $2\theta = 165^\circ$. Multi solution³ and trial and error methods were used, and the parameters were refined by the least-squares method, in the block-diagonal approximation, to a final

Fig. 1 Structure and conformation of the molecule. Note the internal C–H...O hydrogen bonding.

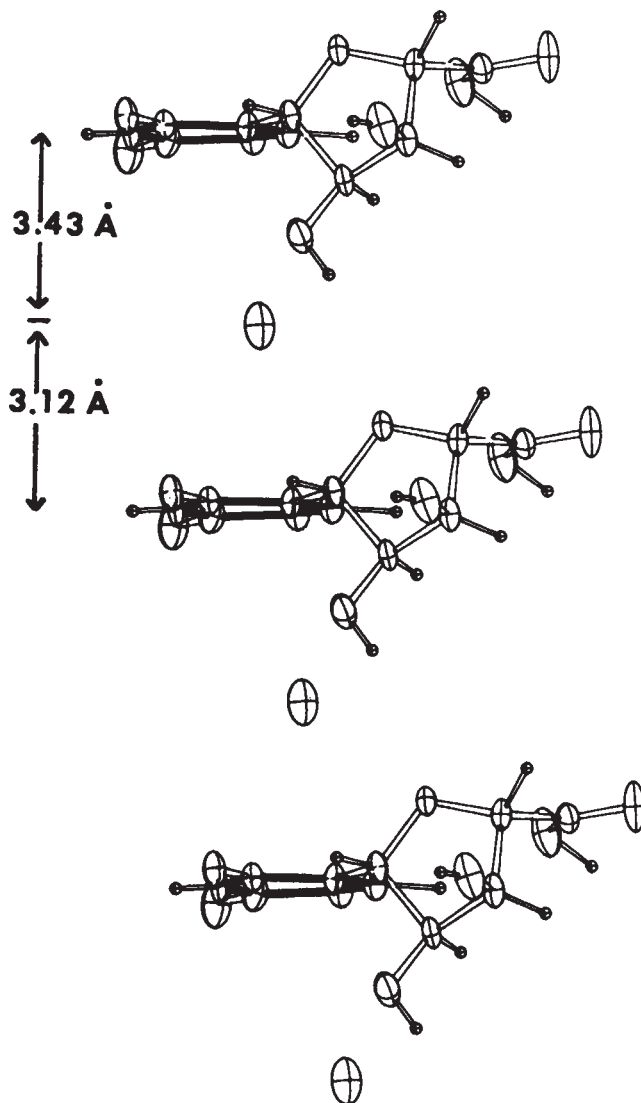
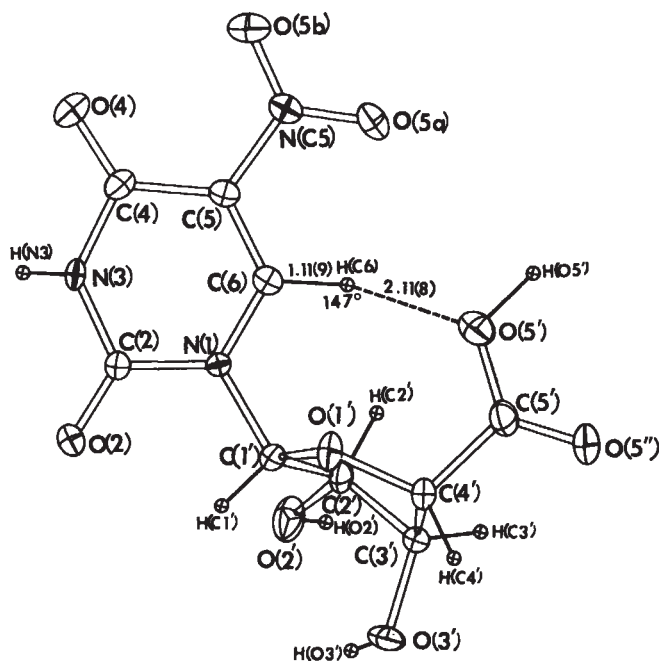


Fig. 2 The 'sandwiching' of water molecules between nucleic acid bases, as seen edgewise. The water molecule here acts as a 'spacer' keeping the bases 6.55 Å apart. The 5-nitro group has been omitted in this figure for clarity.

discrepancy factor of 0.056 for the 1,427 reflections with values of I greater than 2σ . All the hydrogen atoms, except the pair on the water molecule, were located from electron-density difference maps and their parameters were refined using individual isotropic thermal parameters.

The molecular structure and conformation are displayed in Fig. 1. The pyrimidine ring is slightly non-planar with N(1) and C(4) deviating from the least-squares plane through the other four atoms by 0.037 and 0.022 Å, respectively. The nitro group is slightly twisted away from the plane of the uracil ring; the C(4)–C(5)–N(5)–O(5b) torsion angle is 7.6° . The molecule is in the *anti* conformation ($\chi_{CN} = 53.9^\circ$), if we may extend the nomenclature for nucleosides⁴ to this molecule. The furanose ring has the C(2')–endo–C(3')–exo(^{2T}₃) conformation, with C(2') and C(3') deviating from the plane of the other three atoms by 0.418 and 0.199 Å, respectively. The pseudorotation parameters⁵ for the five-membered ring are $P = 171^\circ$ and $\tau_m = 38^\circ$. The torsion angles O(1')–C(4')–C(5')–O(5') and C(3')–C(4')–C(5')–O(5') which describe the relative orientation of the hydroxyl with respect to O(1') and C(3') are 20.4° and 98.3° , respectively. Similar conformation of a β -D-uronic acid has been observed⁶ in thymidine 5'-carboxylic acid. In the usual nucleoside nomenclature, this conformation may be referred to as *gauche*⁺ (*g*⁺, or *gg*). But the conformational preferences of the CH_2OH and $COOH$ groups are different and the torsion angles

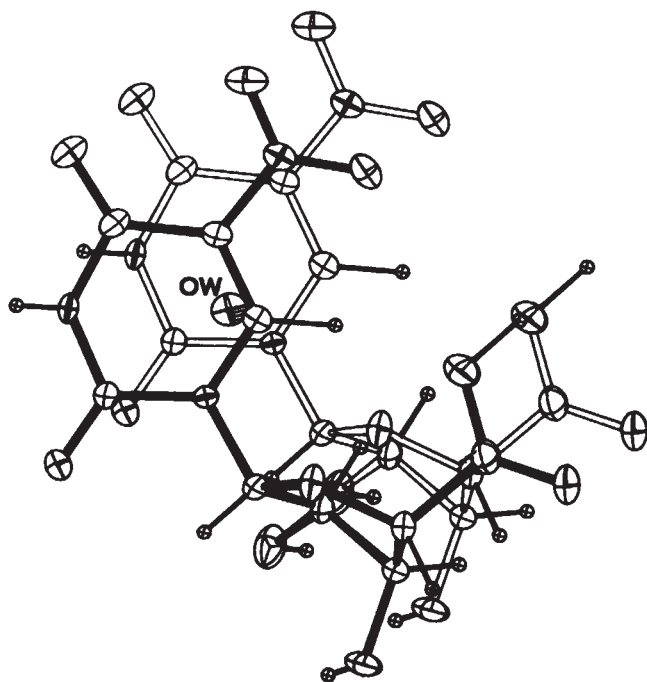


Fig. 3 Partial overlap of the nucleic acid bases and the water in between, viewed normal to the plane of the uracil ring.

are quite different from $\pm 60^\circ$. A more appropriate notation for describing this conformation may be that of Klyne and Prelog⁷; O(5') is (—) *syn*-periplanar to O(1') and (—) *anti*-clinal to C(3').

The *anti* conformation of this molecule is stabilised by an internal C(6)—H(C6) ... O(5') hydrogen bond of length 3.104 Å (Fig. 1). The role of C—H ... O (ref. 8) in conferring some stability to the *anti* conformation of nucleosides and nucleotides has been pointed out by several workers^{9–11}. It is known from deuterium exchange experiments¹² and proton magnetic resonance studies¹³ in solution that the hydrogen on C(8) of purines and C(6) of pyrimidines is acidic. Our results indicate that the location of an 'electron withdrawing' nitro group adjacent to C(6)—H causes additional polarisation of the C—H bond, and thus enables the hydrogen to take part in a C(6)—H(C6) ... O(5') hydrogen bond. This polarisation effect of nitro groups in pyrimidines is analogous to that of a positive charge adjacent to C—H bonds in purines¹, as discussed by us recently. Sundaralingam^{4,10} has pointed out that nucleotides are less flexible than nucleosides. A part of this additional stability in the *anti* conformation of 5'-phosphorylated nucleotides, at least for those nucleotides that can take up a zwitterionic form, is probably due to the charge on the heterocyclic ring and the polarisation of the C(6)—H or C(8)—H bonds.

An interesting feature in the present structure is the wealth of C—H ... O hydrogen bonds and weaker interactions. All carbons in the furanose ring except C(1') take part in C—H ... O contacts in which the H ... O distances and the C—H ... O angles range from 2.27 to 2.41 Å and 130° to 163° , respectively. The uracil rings are not self-paired; N(3) is hydrogen bonded to O(5'') of a neighbouring molecule (N(3)—H(3) ... O(5''), 2.855 Å; H(3) ... O(5''), 2.13 Å; N(3)—H(3) ... O(5''), 168.4°).

The most remarkable feature in this structure is the location of the water molecules; they are between successive, parallel pyrimidine bases of a stack, the members of which are related by translations along *c* and are 6.55 Å apart (Fig. 2). The water molecule is 3.43 Å from the plane of the base above and 3.12 Å from the plane of the base below, thereby forming a 'sandwich' (Fig. 3). The distances are similar to base stacking distances in nucleic acid polymers and crystals of monomers. The water

molecule is held in this location by a strong but bent hydrogen bond (O(5')—H(O5') ... OW, 2.589 Å; H(O5') ... OW, 1.71 Å; O(5')—H(O5') ... OW, 128°) from the carboxyl group of a neighbouring molecule in an adjacent stack. The shortest contacts of the water molecule are with the atoms in the pyrimidine ring below and are: O(W) ... N(1), 3.229 Å; O(W) ... C(2), 3.260 Å; O(W) ... O(1'), 3.305 Å. All other contacts are longer than 3.4 Å. Since the hydrogen atoms of the water molecule cannot be located in electron density difference maps they must be disordered. Thus, it is not possible to decide whether the O(W) ... N(1) contact is a van der Waals' contact or a weak hydrogen bond, but the stereochemistry is unfavourable for the latter.

The stacking of bases does not take place in organic solvents, but does in water; this interaction between bases has been interpreted as 'hydrophobic stacking interactions'¹⁴. It is well known that several heterocyclic dyes and polycyclic hydrocarbons bind to the bases in accord with the intercalation mechanism of Lerman¹⁵. An extensive analysis of the stacking in the solid state¹⁶ indicates that the stacking pattern in a crystal is specific and the bases stack with electronegative heteroatoms forming close contacts with adjacent aromatic systems. In view of these results, the water 'sandwich' we have found is very unusual and, as far as we know, has not been observed before. It is not clear what forces are responsible for the stability of these water-intercalated nucleic acid base stacks. In any case, our results show that such an intercalation of water molecules between nucleic acid bases is possible in certain circumstances. If so, the consequences of water intercalation in nucleic acids should be extremely important and very diverse. We shall mention only one possible example of these consequences. When there is non-complementary opposition—due either to point mutation¹⁷, or to the transient, non-complementary formation of short helices, or to enzymatic excision followed by repair of strands—the non-complementary bases may swing out and assume extrahelical conformations. In such cases, a water molecule could intercalate and act as a 'spacer', giving some stability to the nucleic acid stacks. Such intercalation of water must be very rare in view of the very limited number of mistakes that occur in nucleic acid replication. On the other hand, as our crystal structure demonstrates, such intercalation of water between nucleic acid bases cannot be ruled out. Further work is necessary to discover the conditions favourable to the formation of such hydrated polymers.

We thank Dr D. Harker for discussions, the NCI for financial support and Mrs G. Hazel for technical assistance.

T. SRIKRISHNAN
R. PARTHASARATHY*

Center for Crystallographic Research
Roswell Park Memorial Institute
Buffalo, New York 14263

Received July 20; accepted September 27, 1976.

*To whom correspondence should be addressed.

- Parthasarathy, R., Soriano-Garcia, M., and Chheda, G. B., *Nature*, **260**, 807–808 (1976).
- Wempfen, I., Doerr, I. L., Kaplan, L., and Fox, J. J., *J. Am. chem. Soc.*, **82**, 1624–1629 (1960).
- Germain, G., Main, P., and Woolfson, M. M., *Acta Crystallogr.*, **A27**, 368–376 (1971).
- Sundaralingam, M., in *Conformation of Biological Molecules and Polymers*, 417–455 (Israel Academy of Sciences and Humanities, Jerusalem, 1973).
- Altona, C., and Sundaralingam, M., *J. Am. chem. Soc.*, **94**, 8205–8212 (1972).
- Suck, D., Saenger, W., and Rhode, W., *Biochem. biophys. Acta*, **361**, 1–10 (1974).
- Klyne, W., and Prelog, V., *Experientia*, **16**, 521–23 (1960).
- Sutor, D. J., *J. chem. Soc.*, 1105–1110 (1963).
- Sheffer, E., and Trueblood, K. N., *Acta Crystallogr.*, **18**, 1067–1077 (1965).
- Rubin, J., Brennan, T., and Sundaralingam, M., *Science*, **174**, 1020–1022 (1971).
- Seeman, N. C., Sussman, J. L., Berman, H. M., and Kim, S. H., *Nature new Biol.*, **233**, 90–92 (1971).
- Schweizer, M. P., Chan, S. I., Helmkamp, G. K., and Ts'o, P. O. P., *J. Am. chem. Soc.*, **86**, 696–700 (1964).
- Ts'o, P. O. P., Kondo, N. S., Schweizer, M. P., and Hollis, D. P., *Biochemistry*, **8**, 997–1029 (1969).
- Ts'o, P. O. P., Melvin, I. S., and Olson, A. C., *J. Am. chem. Soc.*, **85**, 1289–1296 (1963).
- Lerman, L. S., *J. molec. Biol.*, **3**, 18–30 (1961).
- Bugg, C. E., Thomas, J. M., Sundaralingam, M., and Rao, S. T., *Biopolymers*, **10**, 175–219 (1971).
- Fresco, J. R., and Alberts, B. N., *Proc. natn. Acad. Sci. U.S.A.*, **46**, 311–321 (1960.)

matters arising

Stability of Lotka–Volterra systems

IN our recent paper in this journal, we claimed that global stability was a consequence of the negative-definiteness of a matrix $B = (A + A')/2$. The subsequent argument was accurate to the extent that if B is found to be negative-definite, then the system must be locally and globally stable. Since A is taken to be an arbitrary matrix, however, it is possible that A can be stable while B is not negative-definite. This means that there may be some locally stable systems that are not globally stable.

The arguments and conclusions are therefore modified as follows. Let d_i be a set of n strictly positive numbers. Introduce the Liapunov function

$$V(x) = \sum_i 2d_i[\exp(x_i) - x_i - 1]$$

which has time derivative $dV/dt = u'(DA + A'D)u$ where u is the vector with elements $[\exp(x_i) - 1]$, and D is a matrix with diagonal elements d_i and zeros elsewhere. This allows us to state the following theorem: if there exists a strictly positive diagonal matrix D such that $(DA + A'D)$ is negative-definite, the Lotka–Volterra system is simultaneously locally and globally stable.

It is not possible to say that all locally stable states are globally stable. The criterion of the previous paragraph is in general a sufficiency condition, not a necessity condition. The criterion does identify the largest known subset of locally stable states that are also globally stable. This is the strongest result available on global stability so far.

Some interesting, if enigmatic, features are worth noting. First, setting $D = I$, the unit matrix, our theorem shows that negative-definiteness of the symmetric part of the community matrix A is sufficient to ensure local and global stability. Second, if $D = N^*$, negative-definiteness of the symmetric part of the interaction matrix α which has elements a_{ij} is also sufficient to ensure local and global stability. Third, the theorem stated above identifies a subset of the possible matrices A which is also D -stable². This means that for this subset stability depends only on the interaction matrix α and is independent of the precise value of N^* , so long as all N_i^* are strictly positive. Such an identification

is in accord with intuitive feelings about the relationship between the geometry of the isoclines of the Lotka–Volterra systems and the flow in phase space.

Finally, there is a strong connection with qualitative (or sign) stability³, which involves the analysis of interaction matrices α where only the signs of the interaction elements are known. Our theorem above shows directly that Lotka–Volterra models which have sign stable interaction matrices α with strictly negative diagonal elements have globally stable dynamics.

Since submitting this manuscript the error in our original paper has independently been pointed out to us by Dr. H. I. Freedman, Department of Mathematics, University of Alberta. In addition Drs. Granero-Porati and Zanaua, Institute of Physics, Parma questioned our use of a community matrix with elements $(a_{ij}N_i^*)$; however we pointed out that such matrix has the same eigen values as a matrix with elements $(N_i^*a_{ij})$.

S. D. TULJAPURKAR

Department of Physics,
Portland State University,
Portland, Oregon 97207

¹ Tuljapurkar, S. D., and Semura, J. S., *Nature*, **257**, 388–389 (1975).

² Barnett, S., and Storey, C., *Matrix Methods in Stability Theory* (Barnes and Noble, New York, 1970).

³ Quirk, J., and Ruppert, R., *Rev. Econ. Studies*, **32**, 311–326 (1965).

Relatedness

ALTHOUGH the article by Dawkins and Carlisle¹ points out relevant questions regarding desertion, it also perpetuates a misconception concerning relatedness. In their example, a female with enough food to support only one infant must choose between two infants the same age, an orphaned baby brother and her own son. They conclude "Intuition points to the son but this is not necessarily correct. There are no genetic grounds for preferring either infant: the mother's relatedness to both is the same, 0.5." There is, however, a genetic difference! The degree of relatedness between two siblings is only on the average 0.5, in any particular instance it may be much less or much greater than that. Additional assumptions are required to state that behaviour is based on average relatedness. Although it could hardly be considered to prove anything, the intuitive

response to save the offspring might, in fact, indicate that nature prefers a sure thing (relatedness=0.5 as in the case of the son) to gambling (average relatedness=0.5 as in the case of the siblings). In addition, in the example they used, even the average relatedness was $\neq 0.5$. Assuming a dioecious diploid organism whose parents are not related, opposite sexed siblings are actually somewhat <50% related on average since they must necessarily have one sex determining chromosome which is not in common. Similarly, siblings of the same sex are on average somewhat >50% related.

It should be noted that the validity of their conclusions is not affected by this correction.

P. A. K. GIBSON

Ecology Program,
University of Illinois,
Urbana, Illinois 61801

¹ Dawkins, R., and Carlisle, T. R., *Nature*, **262**, 131–133 (1976).

DAWKINS REPLIES—Gibson¹ is right, both in the main point she makes and in her observation that it doesn't matter anyway. But her gambling analogy is misleading.

She thinks of the offspring as a safe bet while the sibling is a risky gamble. Yes, setting aside Gibson's trivial point about sex chromosomes, it is true that exactly 50% of the genome of a parent is inherited by a given child, while siblings share 50% of their genomes only on average. But who cares about genomes? Certainly not natural selection, at least in sexual populations². From the point of view of a single altruistic gene sitting in the body of an individual, a particular child of that individual is just as risky a prospect as a particular sibling. The gene may or may not be present in the body of the offspring, and it may or may not be present in the body of the sibling. It is a 50% gamble in both cases.

Department of Zoology,
University of Oxford, UK

¹ Gibson, P. A. K., *Nature*, **264**, 381 (1976).

² Dawkins, R., *The Selfish Gene* (Oxford University Press, 1976).

Fast axonal transport and amine levels

THE monoamine oxidase inhibitor, pargyline, which raises amine levels, produces a myopathy which can be prevented by previous nerve section¹. The fact that the pargyline-induced myopathy could be prevented by nerve section led Boegman and collaborators to look at effects of this drug on axonal transport. They reported that pargyline treatment caused a dramatic increase in the velocity of axonal transport², and so I decided to see what effects the reported increased velocity of transport might have on life span or composition of transported protein.

Male Sprague-Dawley rats (250–300 g) were injected intraperitoneally with the following solutions: (1) 1 ml saline for 3 d before the experiment (controls); (2) 25 mg kg⁻¹ pargyline (Parsidol, Abbott Pharmaceuticals) in 1 ml saline for 3 d; (3) 75 mg kg⁻¹ pargyline for 3 d; (4) 2.5 mg kg⁻¹ reserpine (Serpasil, Ciba) for 7 d. Injection of ³H-L-leucine into the ventral horn of the lumbosacral cord was performed as described previously³. At intervals after injection of precursor, the sciatic nerve and ventral roots were dissected out, cut into segments, and placed in 10% TCA overnight. TCA-insoluble material was dissolved in Protosol (New England Nuclear) and counted. In a few experiments TCA-soluble activity was counted by adding Scintiverse (Fisher). In some animals whole brain, right atrium and soleus muscles were dissected out and processed according to the Falck-Hillarp technique; in others, whole brains were assayed fluorimetrically for noradrenaline (NA) and dopamine (DA).

Drug treatment was judged to be effective: treated animals showed characteristic changes in appearance and behaviour^{1,4}. There was complete loss of amine-specific fluorescence in reserpine-treated animals, and an increase in fluorescence intensity of aminergic neurones in pargyline-treated animals. Pargyline produced significant increases of 224% in whole-brain NA and 167% in whole-brain DA.

After injection of ³H-leucine, a wave of labelled protein moved along the axons of the sciatic nerve. The velocity of transport was given by the slope of the regression line on a plot of wavefront position against elapsed time. Neither reserpine nor pargyline treatment had any effect on the velocity of transport (Fig. 1). To eliminate the possibility that some labelled protein was transported at higher velocities ahead of the main wavefront, a second series of experiments was carried out.

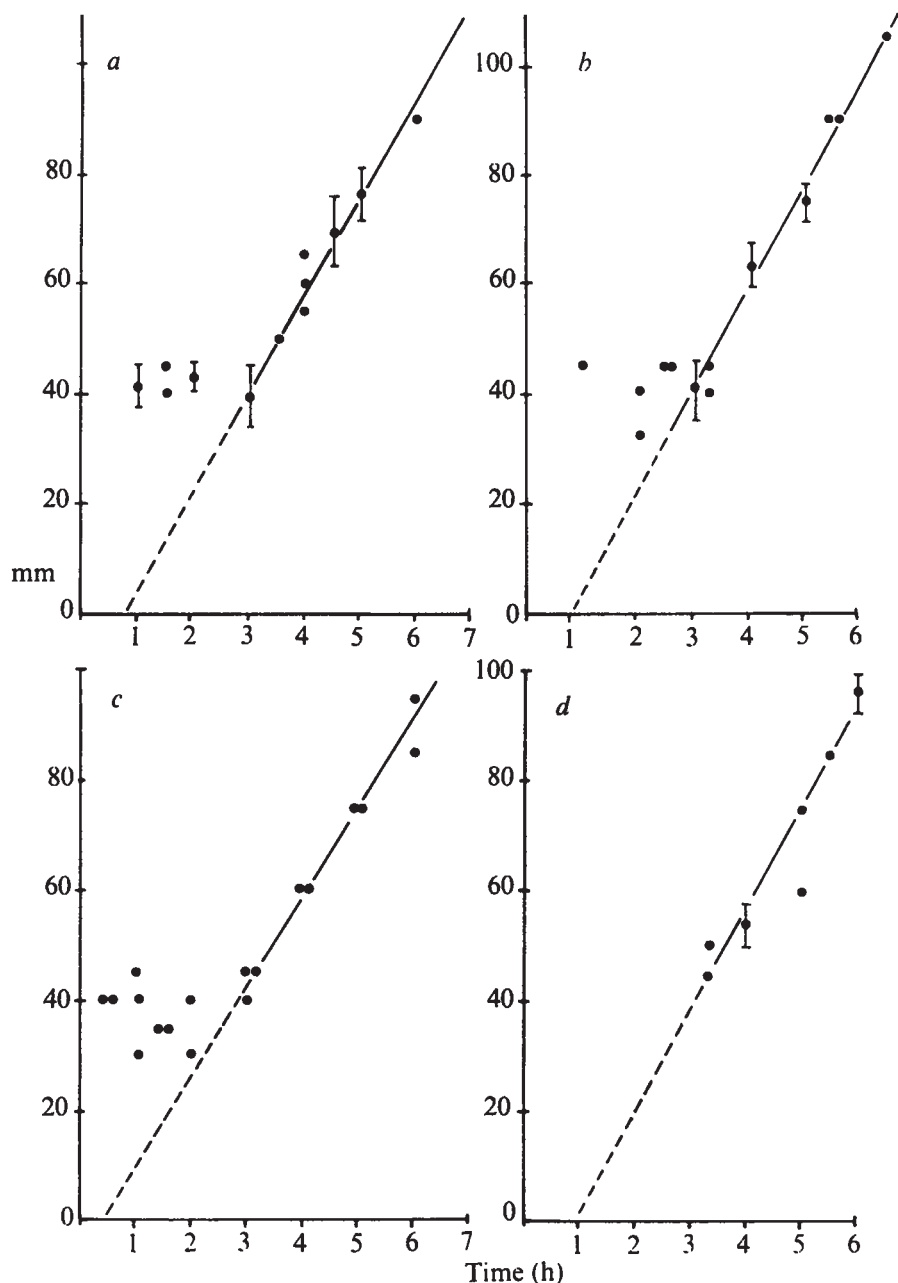


Fig. 1 Velocity of axonal transport. Ordinates, distance of wavefront of activity from most proximal site of injection; abscissae, time elapsed since injection. Where 4 or more points were obtained for any time, the means $\pm$ s.e.m. have been plotted. Regression lines were fitted by the least squares method and were extrapolated to the time axis. Note that at intervals of less than 3 h the wavefront remained stationary at approximately 40 mm from the site of injection. Velocities of transport calculated from the slope of the regression line were: a, controls, 431 ± 16 mm d⁻¹ (mean $\pm$ s.e.m.); b, pargyline (75 mg kg⁻¹), 458 ± 26 ; c, pargyline (25 mg kg⁻¹), 391 ± 69 ; d, reserpine (2.5 mg kg⁻¹) 448 ± 35 .

In a group of rats treated with pargyline (25 mg kg⁻¹) the sciatic nerve was crushed at the time of precursor injection, and the time when transported protein began to accumulate proximal to the crush was measured. The velocity of transport, calculated from the onset of accumulation, was 412–436 mm d⁻¹.

At distances of less than 40 mm from the site of injection, the velocity of axonal transport could not be measured reliably owing to the presence in the

ventral roots of high ³H activity (Fig. 1). This activity appeared within 30 min of the precursor injection. Three hours after injection the 'mobile' wavefront broke away from this 'stationary' wavefront and proceeded along the axon. In the region of the stationary wavefront, 50% of the total ³H activity was TCA soluble and presumably represented the diffusion of free ³H-leucine from the site of injection. Beyond the 40-mm limit, which coincided with the exit of the sciatic nerve roots from the

spinal canal, the TCA-soluble activity fell to 10%. Boegman *et al.*² measured total ³H activity and obtained wavefronts located 40 mm from the injection site: there is no indication that they followed wavefronts into more distant regions of the sciatic nerve in order to obtain regression lines for determination of velocity. I suggest that the apparent increase in velocity that they report results from measurements made on a stationary wavefront of diffused activity at shorter time intervals for pargyline-treated than for normal animals.

It therefore seems that changes in amine levels produced by pargyline and reserpine have no effect on the velocity of axonal transport of protein. A further study³ of axonal transport involving manipulation of amine levels requires re-evaluation in the light of these results.

The work in my laboratory is supported by the Medical Research Council of Canada. I am grateful to Ms G. Bilsborough and Ms B. Hatf for assistance and to Mr L. Bause for the fluorimetric assays. Pargyline (Parsidol) was provided by Abbott Pharmaceuticals.

M. A. BISBY

*Division of Medical Physiology,
Faculty of Medicine,
University of Calgary,
Calgary, Alberta T2N 1N4, Canada*

¹ Yu, M. U., Wright, T. L., Dettbarn, W. D., and Olson, W. U., *Neurology*, **24**, 237-244 (1974).

² Boegman, R. J., Wood, P. L., and Pinaud, L., *Nature*, **253**, 51-52 (1975).

³ Bisby, M. A., *Expl Neurol.*, **50**, 628-640 (1976).

⁴ Steg, G., *Acta Physiol. scand.*, **61**, suppl. 225 (1964).

⁵ Boegman, R. J., and Wood, P. L., *J. Neurochem.*, **26**, 737-740 (1976).

BOEGMAN *et al.* REPLY—We agree with Bisby¹ that pargyline does not alter the bulk flow of material down the rat sciatic nerve. In our published data² we reported on the total radioactivity present in nerve segments and not the TCA-precipitable activity. This led us to conclude that pargyline alters the rate of fast axoplasmic flow.

We disagree with Bisby's comment³ that axonal flow rates cannot be measured within the first 40 mm since we have shown that puromycin will abolish a front of TCA-precipitable radioactive material within this distance.

*Department of Pharmacology,
Queen's University, Kingston,
Canada*

¹ Bisby, M. A., *Nature*, **263**, 382-383 (1976).

² Boegman, R. J., Wood, P. L., and Pinaud, L., *Nature*, **253**, 51-52 (1975).

Large scale extinctions

THOMSON'S contribution¹ to the search for the cause of major extinctions points out some patterns of temporal

diversity change that merit consideration. I wish to note an important inconsistency in Thomson's argument and to discuss some work which contradicts his hypothesis of extinction through specialisation.

Thomson mentions that the diversification portion of his diversity curves are typically exponential, not logistic in shape. Later, however, he describes this diversification as "species packing" and consequent niche reduction". Such species packing should result in a logistic growth of diversity², not the exponential one that he notes. Cisne³ demonstrated such a logistic increase in the morphological specialisation of Phanerozoic aquatic free-living arthropods and also showed that diversification accompanies this specialisation, as suggested by Thomson. The apparently broad time intervals used by Thomson may prevent the detection of a logistic pattern of diversification.

That specialised taxa tend to become extinct is one of the most plausible assumptions of evolutionary biology and forms the basis of Thomson's explanation. The only test of the extinction through specialisation hypothesis failed, however, to detect any significant correlation between specialisation and evolutionary longevity within aquatic free-living arthropods³. Furthermore, Thomson's suggestion that diversification (and specialisation) is "entrained" may inadvertently revitalise the corpse of orthogenesis.

Thomson has pointed out some phenomena of major evolutionary significance—the apparent symmetry of the diversity curves and the existence of "group diversity curves"^{4,5}. I agree that it is important to focus on the causes of these phenomena rather than on one striking but merely consequential aspect—extinction.

KARL W. FLESSA

*Department of Earth and Space
Sciences,
State University of New York at
Stony Brook,
Stony Brook, New York 11794*

¹ Thomson, K. S., *Nature*, **261**, 578-580 (1976).

² Strong, D. R., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 2766-2769 (1974).

³ Cisne, J. L., *Evolution*, **22**, 337-366 (1975).

⁴ Flessa, K. W., Powers, K. V., and Cisne, J. L., *Paleobiology*, **1**, 71-81 (1975).

⁵ Newell, N. D., *J. Paleontol.*, **26**, 371-385 (1952).

⁶ Flessa, K. W., and Imbrie, J., in *Implications of Continental Drift to the Earth Sciences*, 1 (edit. by Tarling, D. H., and Runcorn, S. K.), 247-285 (Academic, London and New York, 1973).

THOMSON REPLIES—Flessa¹ has read into my highly qualified statements more than I actually said. I merely stated that the diversity curves "approached" an exponential pattern. I tried to explain the curves only by "analogy" with "species packing" and

in any case was working at the genetic level. I do not believe that the patterns of diversification that I described have anything to do with orthogenesis. The question of how one would recognise in fossils the specialisations that could result in extinctions is a very difficult one. Not all specialisations will necessarily be associated with extinction and the reverse is also true. If analogy with any process like species packing can be upheld, it will probably turn out that the crucial specialisations were ones of behaviour and ecology rather than of major external morphology. In that case one might expect morphological specialisation to correlate better with diversification than extinction.

*Peabody Museum of Natural History,
Yale University*

¹ Flessa, K. W., *Nature*, **264**, 383 (1976).

The creeping-film phenomenon of potassium chloride solution

BIN-JUINE HUANG and JEN-CHI HUANG¹ state that the creeping-film phenomenon of potassium chloride solutions cannot be satisfactorily interpreted "based on present knowledge of either surface chemistry or transport theory." We believe the phenomenon can be explained by using ordinary physico-chemical surface forces such as capillary action in a type of 'wick' effect.

Consider a glass beaker containing a saturated solution of potassium chloride; when some of the solution evaporates, crystals form on the sides of the beaker at the surface of the solution. The solution is then drawn up into micro-fissures and crevices in the crystals by capillary action, and when the solution reaches the edge of the crystal formation it wets the glass surface, and on further evaporation forms more crystals which advances the crystal front causing the creeping phenomenon observed. This is analogous to a 'wick' effect where the KCl crystals act as a wick for the solution. This would explain the creeping which occurs over the brim of the beaker.

A saturated KCl solution will exhibit creeping effects when it wets a substrate surface, and this has been observed in a variety of substances, such as glass, aluminium and lucite plastic (A. M. Yacynych, unpublished). It would be difficult to explain this phenomenon by invoking a specific interaction between the solution ions and the substrate surface, in view of the diverse chemical nature of the substances on which this phenomenon occurs. This creeping effect does not occur when the KCl solution does not wet a substrate surface, as is the case with Teflon (A. M. Yacynych, unpub-

lished). Also, creeping does not occur if the KCl solution is unsaturated, or if a beaker containing a saturated solution is stoppered. In the case of an unsaturated solution, if any crystals were formed on the sides of the beaker they would be dissolved as the unsaturated solution was drawn into any fissures or crevices in the crystal and the creeping could not occur. In the case of a covered beaker, no creeping occurs because the atmosphere inside the beaker is saturated with water vapour, preventing evaporation of the solution which is necessary for the crystal front to form and advance.

Regarding the creeping which occurs through a 'seamless' glass tube connector; the authors do not specifically mention whether the joint was greased or ungreaed. In the former case, we would not expect any creeping to occur, however, in the later case, small fissures and crevices are certain to be present in the joint allowing the creeping to occur. In fact, this ability of the KCl solution to creep into small fissures is put to good use by electrochemists in constructing reference electrodes². Instead of using a porous frit or a salt bridge to maintain solution contact, a platinum wire is sealed in glass and contact is maintained by way of the small fissures which form between the platinum and glass on cooling.

The alkali halides form two separate isomorphous groups³: the lithium and sodium salts; and the potassium, rubidium, and caesium salts. If the creeping results simply from the wick effect by the crystals, one would not expect it to be unique, and in fact, compounds such as potassium bromide, potassium iodide, and caesium chloride that are crystallographically isomorphous with KCl also exhibit this creeping phenomenon (A. M. Yacynych, unpublished). Sodium fluoride, chloride, bromide and iodide which are not isomorphous with KCl do not creep. Perhaps their crystal formation does not have the small fissures and crevices necessary for the wick action.

ALEXANDER M. YACYNCH
WILLIAM R. HEINEMAN

Department of Chemistry,
University of Cincinnati,
Cincinnati, Ohio 45221

CONCERNING the paper by Huang and Huang¹, may I suggest that the well known Marangoni Effect may provide the authors with an explanation for their observations. A good account of different types of Marangoni Effect may be found in ref. 2.

D. A. J. RAND

CSIRO,
Minerals Research Laboratories,
Division of Mineral Chemistry,
PO Box 124, Port Melbourne,
Victoria, Australia 3207

¹ Huang, B. J., and Huang, J. C., *Nature*, **261**, 36-38 (1976).

² Scriven, L. E., and Sternling, C. V., *Nature*, **187**, 186-188 (1960).

Matters arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

HUANG and Huang¹ describe the flow of potassium chloride solutions over surfaces and conclude by saying the results are "inexplicable" but that they might be attributed to "interactions of the ions with the surface of the beaker or tube and the surrounding gaseous molecules". In 1950 I observed the same effect in aqueous solutions of potassium ferricyanide, but results of several experiments were not published because I was unable to obtain reproducible quantitative results. I attributed the phenomenon to surface mobility of ions and molecules.

Evidence for surface mobility of adsorbed atoms was first given by Volmer and Estermann in 1921. They set up an experiment in which mercury evaporated from a bath at -10°C to a glass plate at -65°C . Hexagonal crystal plates of mercury formed on the glass. The crystals should have grown at the rate G/ρ in which G is the rate at which mercury molecules struck unit area of crystal at the glass surface and ρ is the density of the crystal. But the dimensions of the crystals grew $1,000\times$ faster than that. Volmer and Estermann concluded the

mercury atoms moved over the surface of the glass to become attached to the crystals.

A few years later, Volmer and Adhikari placed a benzophenone crystal between 0.02 to 0.09 mm from the edge of a glass plate; a mercury drop just touched the edge of the glass. The crystal dissolved in the mercury due to the mobility of benzophenone molecules over the surface of glass (vapour pressure negligible); it was estimated that the resistance to diffusion over the glass was one hundredth of that in the liquid state.

Wicke in 1940 was the first to suggest that molecules adsorbed from the vapour phase contribute to total flow through porous materials; and the following year Wicke and Kallenbach flowed mixtures of CO_2 and N_2 past one face of a charcoal membrane, at 600 torr total pressure and pure N_2 past the other face at the same pressure. In the absence of a pressure gradient there was no flow through the membrane but diffusion of CO_2 into N_2 could occur and the diffusion coefficient could be determined by analysis. Experiments were done in the range $0-300^{\circ}\text{C}$. From the laws of gas kinetics, the diffusion should increase with temperature and this was observed at high temperatures. At low temperatures, however, the diffusion coefficient increased as temperature was decreased, which was attributed to the greater adsorption of CO_2 at lower temperatures. Mobility in the adsorbed phase contributed as much as 50% of the total flow of CO_2 at the lowest temperature.

So, when crystals are in contact with another solid surface an adsorbed layer of atoms (or molecules) can form on the solid surface in dynamic (mobile) equilibrium with the crystal. When a vapour (or gas) is in equilibrium with a solid surface there will be an adsorbed phase on the solid in kinetic equilibrium with the vapour; atoms or molecules will be mobile over the surface within the adsorbed phase. Evidence for this is considerable²⁻⁴. Adsorption also occurs when liquid mixtures and solutions are in contact with solid surfaces and various forms of adsorption isotherm have been found experimentally, see for example, Kipling⁵. Atoms and molecules will be mobile within that adsorbed layer.

W. M. JONES

Department of Physics,
The University College of Wales,
Penglais,
Aberystwyth, Cardiganshire, UK

¹ Huang, B., and Huang, J., *Nature*, **261**, 36 (1976).

² Brunauer, S., *The Adsorption of Gases and Vapours* (Oxford University Press, 1944).

³ Carman, P. C., *Flow of Gases and Vapours through Porous Media* (Butterworths, London, 1956).

⁴ Roberts, G. T., *J. Phys. A.*, **6**, 570 (1973).

⁵ Kipling, J. J., *Adsorption from Solutions of Nonelectrolytes* (Academic, London and New York, 1965).

¹ Huang, B. J., and Huang, J. C., *Nature*, **261**, 36-38 (1976).

² Heineman, W. R., and Kuwana, T., *Analyt. Chem.*, **44**, 1972-1978 (1972).

³ Mellor, J. W., *A Comprehensive Treatise on Inorganic and Theoretical Chemistry*, **11**, 531-532 (Longmans, Green and Co., New York, 1922).

reviews

Golden jubilee monograph

J. H. S. Blaxter

The Eggs and Planktonic Stages of British Marine Fishes. By Sir Frederick S. Russell. Pp. xv+524. (Academic; London, New York, San Francisco, 1976.) £19.50.

THERE is a tendency, as fisheries come under increasing pressure, for the success of the fishery to become dependent on the recruit fish rather than on a longer run of successful year-classes. The study of the pre-recruit phases of the life history, both for the purpose of prediction as well as for understanding what determines success or failure of a year-class, then assumes greater importance. This is one of the reasons why investigations of eggs, larvae and juveniles of commercial fish species have been greatly stepped up over the past twenty years. Both sea surveys of these young stages, as well as laboratory experiments involving rearing from the egg, are now part of the standard repertoire of most larger fisheries laboratories, and it is now generally accepted that the year-class strength is probably determined in the first year of life.

It is surprising that no standard work of identification of the young stages of north-eastern Atlantic species has been written since Ehrenbaum's *Eier und Larven von Fischen* which formed the first volume of his *Nordisches Plankton*, produced between the years 1905 and 1909. Many imperfections in identification existed at that time. To a limited extent they were made good in subsequent publications on restricted groups of species such as the gadoids by Johannes Schmidt in 1909, or in more comprehensive descriptions for other areas, as in the *Fauna and Flora of the Gulf of Naples* during the 1930s and in the ICES plankton identification sheets that appeared after World War II.

Sir Frederick Russell's book thus fulfills a clear need. It is a first class and completely fresh re-appraisal of the identification problem. It covers the planktonic stages of coastal species of northern European waters likely to be found within the 200-m depth contour but it excludes strictly oceanic species. Introductory chapters on terminology, development, principles of identification, feeding habits

and vertical and horizontal distribution are followed by the main section where each teleost family is taken in turn. A description, with key where appropriate, is given for the egg, larva and post-larval stage. Relevant ecological or experimental work is also referred to—one aspect of the book which makes it quite different from previous publications.

The value of the monograph lies therefore in the collation of the taxonomic and other publications, and the assembling and up-dating of the many keys scattered in the literature. During this process the author has brought to bear his unique experience in this field, to improve, clarify and amplify the keys.

The reference to many 1975 and even some 1976 scientific papers in the bibliography is evidence of his success in making the monograph as comprehensive and up-to-date as possible within the range of species adopted. It

is, however, a pity that the style and content of the book, limited as it is to northern European fishes, mean that some of the excellent and relevant ecological and experimental work done, for example, on the Californian sardine and anchovy, has perforce had to be excluded.

Sir Frederick Russell retired from the directorship of the Plymouth laboratory in 1969 and this monograph is a *tour de force* representing the results of his enviable energy during subsequent years. He published his first paper on fish eggs and larvae in 1926 so that the monograph is a fitting and admirable golden jubilee offering to the scientific community. We all look forward to the celebration of his diamond jubilee! □

Dr Blaxter is a Senior Principal Scientific Officer at Dunstaffnage Marine Research Laboratory, Oban, Argyll, UK.

Energetic ion beams and sputtering

Ion Implantation, Sputtering and Their Applications. By P. D. Townsend, J. C. Kelly and N. E. W. Hartley. Pp. ix+333. (Academic; London and New York, June 1976.) £13; \$32.25.

IN a great many practical situations the behaviour of a component depends on the near-surface composition, physical state or topography of the material. Examples spring to mind in the optical, mechanical, electrical and corrosion fields, and the extraordinary developments in semiconductor devices and circuits involve a sophisticated control of material properties within a layer a few microns thick.

Diffusion, coating and etching have hitherto been used to control surface composition and its lateral variation, but these are no longer precise, versatile or reproducible enough for some requirements. The past decade has seen the development of an entirely different set of techniques based on the use of energetic ion beams by means of which chosen atomic species can be injected into a surface, or, by sputtering, material can be removed.

This book provides a very timely review of the progress achieved so far

in a field in which industrial applications are being extended as each month goes by. The fundamentals of the subject are treated admirably, from the slowing-down of the incident ions to their effects on the target structure. The treatment of applications is more superficial, but the authors must be conscious of the speed with which a detailed account would become outdated. Readers, and there will be many, who are mainly interested in the ion implantation of semiconductors will be disappointed: the literature of this is now so extensive that the authors, perhaps wisely, have chosen to concentrate on the remainder.

A bonus, for which the reader is not well prepared by the title of the book, is a lengthy and most useful chapter dealing with the analytical applications of ion beams.

I congratulate the authors on their clear and readable treatment, and the publisher for the attractive presentation of text and diagrams.

G. Dearnaley

Dr Dearnaley is a member of the research staff at the Atomic Energy Research Establishment, Harwell, UK.

Describing the brain

Cell Biology of the Brain. By W. E. Watson. Pp. xi+527. (Chapman and Hall: London, August 1976.) (Distributed in the USA by Halsted: New York.) £15.

THERE have been previous attempts to describe the brain in the language of cell biology, but few have been as crammed with information or as well written as the present monograph. The author has not been intimidated by this heroic task, and has produced an eminently readable and useful volume. The first six chapters give a broad outline of the cellular properties of the brain, its origin from epithelial tissue, the differentiated properties of neurones, current understanding of membranes and their function, and a fairly extensive review of modern knowledge of the 'chemical transducer' mechanisms by which neurones respond to chemical signals.

This includes a description of recent biochemical techniques for the study of receptor sites by the ligand binding approach, the coupling of receptors to cyclic nucleotide mechanisms and more general biochemical changes in neurones associated with activity. In this section there is little emphasis on the neurone as a carrier of electrical signals, nor is there any description of the presynaptic machinery associated with the production, storage and release of neurotransmitters—a deliberate omission since the author rightly points out that these topics are already dealt with extensively elsewhere. But this perhaps leaves some major gaps in coverage for student readers.

The largest chapters, and the most valuable in my estimation, were the remaining four, which deal with the topics of plasticity in the nervous system, responses to injury, neurotrophism and genetics. These are topics which encompass some of the most fundamental problems of neurobiology, and the author is an expert in this area. He gives an excellent review, taking examples from a broad spectrum of vertebrate and invertebrate studies. In dealing with this large area the author deliberately, and probably wisely, avoids the even larger topic of developmental neurobiology, but his deliberate omission of any discussion of the nerve growth factor seemed to me a curious decision.

The book is undoubtedly well referenced. Indeed some might consider that the 3,247 citations in the bibliography (which occupies almost 200 of the 527 pages) were enough for six

volumes of this size. I would personally have preferred to have seen fewer references and more illustrations. Overall, however, there is no doubt that the book is an unusual and valuable source of information on topics that are not well reviewed elsewhere.

Leslie Iversen

Leslie Iversen is Director of the MRC Neurochemical Pharmacology Unit at the University Department of Pharmacology, Cambridge, UK.

Neuromuscular junction

Neuromuscular Junction. (Handbook of Experimental Pharmacology. New Series, Vol. 42.) Edited by Eleanor Zaimis. Pp. xvii+746. (Springer: Berlin and New York, 1976.) DM 280; \$91.90.

THIS latest volume in this well-known series of reviews and monographs, which all pharmacologists turn to for information when everything else fails, lives fully up to the standards of its predecessors. Each of its chapters amounts to a substantial review article and like others in the series, the book as a whole is full of detail, with each chapter accompanied by an exceptionally complete and valuable bibliography.

The neuromuscular junction is, without any doubt, the most intensively studied synapse of all; it is also a highly atypical synapse, being a simple one-to-one connection between a nerve fibre and a muscle fibre. In life it serves simply as a relay (and a highly reliable relay at that, very rarely affected by disease) and has none of the integrative properties that characterise most other synaptic connections. To some investigators studies of the neuromuscular junction are justified mainly by the implications that the results hold for other more complex synapses. Not to all, though, for at a recent scientific meeting an eminent investigator, asked whether the neuromuscular junction helped to explain brain function, replied that all he could be sure of was that the brain was indispensable in understanding the neuromuscular junction.

This book certainly does not lend much encouragement to the view that the neuromuscular junction is a good model for other synapses. The admirable article by MacIntosh and Collier on the metabolism of acetylcholine makes a point of comparing the results of work on the neuromuscular junction, where electrophysiological

techniques have generally proved much more valuable than biochemical techniques, with work on the brain, where the reverse is true. The differences between the two are a good deal more obvious than the similarities, and the authors comment gloomily: "The extent of these peculiarities of individual tissues... may seem depressing to investigators who would like to think that their results have a general significance and should help to explain all analogous phenomena". Particularly valuable in this article is the balanced discussion of the present status of the vesicle hypothesis. The evidence, which is complicated and often conflicting, is presented with an economy and a clarity which is all too rare in scientific writing.

The following chapter, by Ginsborg and Jenkinson, equally well-written, surveys the mechanism of impulse transmission mainly from an electrophysiological point of view; it, too, succeeds in distilling an immense amount of information into a concise and coherent account. What emerges most clearly from these two chapters is the considerable uncertainty that still exists about the mechanism of acetylcholine release and the difficulty in relating the results of electrophysiological experiments to biochemical events within the nerve terminal. These two chapters, together with the preceding one on morphology by Bowden and Duchon, form the first half of the book and complement each other extremely well.

The rest of the book is concerned mainly with neuromuscular blocking drugs and anticholinesterases. The contentious areas here are the mechanisms of action of depolarising blocking agents and of anticholinesterases, well-scarred battlegrounds on which relatively little new evidence has appeared in recent years. I would have preferred to see less discussion of these topics and fuller accounts of the recent biochemical advances in isolating and characterising acetylcholine receptors. Studies based on these advances have already led to a reappraisal of the nature of the defect in myasthenia gravis and show great promise of leading to a much better understanding of the process underlying synaptogenesis and denervation effects; yet they receive only fleeting comment. The book will, however, be invaluable as a work of reference. The only danger is that it may attract even more experimenters to assault this unoffending synapse.

H. P. Rang

H. P. Rang is Professor of Pharmacology at St George's Hospital Medical School, London, UK.

Extragalactic astronomy

Galaxies and the Universe. (Stars and Stellar Systems: Compendium of Astronomy and Astrophysics, Vol. 9.) Edited by A. Sandage, M. Sandage and Jerome Kristian. Pp. xxii+818. (University of Chicago: Chicago and London, 1976.) £30; \$45.00.

It is 50 years since Edwin Hubble discovered Cepheid variable stars in the M31 galaxy, an epochal step in surveying the extragalactic universe, 21 years since this series was conceived, and 14 years since Allan Sandage commenced the editorial work for this book. The result is a book which is a product of the 1960s and which contains articles completed between 1965 and 1973, with a mean epoch of 1970. Is a book six years out-of-date worth £30?

Fortunately that is an improper question: the 19 chapters are written by people with the highest reputation, each contributing in a field in which he or she is an expert. As Sandage rightly says in his preface, one remarkable feature is the enduring quality of the chapters, which are definitely not out-of-date, although some (for example, quasars, clusters, radio emission) are certainly not up-to-date. The editorial brief was to provide a compendium of essential facts about the Universe, and this policy has ensured that ephemeral theoretical playthings and tentative observations of implausibilities have been kept to the irreducible minimum. Not an easy task in extragalactic astronomy!

Galaxies and the Universe provides a review of well-established knowledge on the observed properties of galaxies and quasars. Individual chapters discuss the contents, masses, magnitudes, colours, sizes, energy distributions, and radio properties of galaxies and quasars. There are also extensive contributions on the dynamics and structure of galaxies. Observational cosmology is dealt with in sections on the distance scale of the universe, the distribution of galaxies, and an impressive review by the editor entitled "The Redshift".

The style is that of a review addressed to an informed reader, and there are very extensive bibliographies. Over 1,000 astronomers feature in the Name Index, and 1,200 objects in the Galaxy Index: these statistics give an impression of the ambitious scope of the work.

Galaxies and the Universe stands as a fitting synoptic monument to a heroic age in the quest to explore the Universe. Here are the results of a golden half-century that successively established the existence of island galaxies

beyond the Milky Way, the redshift, radio galaxies, the cosmic background radiation, and quasi-stellar objects. The nature and history of the Universe on the largest scales is now yielding to the methods of science, and the main thrust of astronomical research for the rest of this century will be among the galaxies. Here in broad outline is a competent summary of what we know about those remote objects, and an indication of what we need to find out. It is an absolutely invaluable work of reference for anyone working in the field and an essential purchase for libraries. It is probably the most significant astronomical book published this year.

Simon Mitton

Simon Mitton is Secretary of the Institute of Astronomy, University of Cambridge, and Editor of the Quarterly Journal of the Royal Astronomical Society.

Amphibian physiology

Physiology of the Amphibia. Vol. 3. Edited by Brian Lofts. Pp. xiv+644. (Academic: New York and London, July 1976.) \$58.50; £35.70.

BECAUSE of their phylogenetic position and varying degrees of adaptation to life in two media the Amphibia are a group with which comparative physiologists are rightly very much concerned. In this, the presumably final, volume of *Physiology of the Amphibia*, the somewhat scattered literature in several fields is brought together and critically reviewed. The completed series now constitutes a source of information which should be indispensable to all herpetologists who are physiologically inclined. The timescale for the appearance of the various parts has, however, been unduly extended and it is a measure of the advances which have been made in some aspects of the subject in the intervening twelve years that there are articles on metamorphosis in both volumes one and three. Regrettably the prices of these same two volumes are an even more striking testament to the pace of inflation.

The present volume covers integrative aspects such as colour change (J. T. Bagnara), moulting (L. L. Larsen) and resistance to desiccation (E. Elkan); cell culture (K. A. Rafferty) and immunity mechanisms (E. L. Cooper); general neuroanatomy (A. Oksche and M. Ueck); more specialised articles on the visual (D. Ingle) and auditory (R. C. Capranica) systems; and finally metamorphosis, with special emphasis on its biological and eco-

logical aspects (M. H. I. Dodd and J. M. Dodd). The declared editorial policy has been to group related topics in the same volume but perhaps inevitably the final one seems to be the most heterogeneous, although there is more interrelation between some of the chapters than might at first appear. For instance it is helpful to have the account of the neuroendocrine systems alongside the discussion of the role of neurosecretion in metamorphosis.

The authors of the articles on cell culture and immunity mechanisms both stress the advantages of using amphibian material for general work in these fields; if this should lead to stocks of Amphibia being maintained in more laboratories, then the homily on amphibian husbandry in the chapter on pathology (E. Elkan) is very timely. At least as far as the more terrestrial species are concerned it must be realised that Amphibia in captivity are pathological specimens unless the conditions closely reproduce those found in the natural habitat.

The book is well produced and printing errors are rare, and then only of a minor nature.

C. L. Smith

Dr Smith is a Lecturer in the Department of Zoology at the University of Liverpool, UK.

Are you involved in cancer research?

Do you read

cancer letters?

Co-managing editors
Dr. P. Shubik
and Dr. D. Clayson

The international journal which specialises in the rapid publication of short reports on all aspects of cancer research.

For a subscription and/or a specimen copy, please write to the Promotion Manager,
ELSEVIER/NORTH-HOLLAND BIOMEDICAL PRESS B.V.,
P.O. Box 1527, Amsterdam,
The Netherlands.

0399

Circle No. 20 on Reader Enquiry Form.

Immunobiology of reproduction

The Immunobiology of Mammalian Reproduction. By A. E. Beer and R. E. Billingham. Pp. xvi+240. (Prentice-Hall: Englewood Cliffs, New Jersey, 1976.)

In their preface to this volume, Beer and Billingham point out the absence of any previous concise, but comprehensive, account of the principles of the immunobiology of reproduction. There is of course good reason for this deficiency. A problem which requires considerable knowledge and understanding of subjects as diverse as immunology, oncology, reproductive and developmental biology, genetics and endocrinology makes discussion hazardous if the authors are not to be caught in the crossfire of experts. That this demanding combination of talent and expertise is required is testimony to the fundamental importance to biology of the problems under consideration. If we are to judge this ambitious work by the authors' professed aim to make good the existing deficiency, we can consider the volume as at least a qualified success.

Not surprisingly, the authors cover comprehensively that data on immunological aspects of foeto-maternal relations available up to mid-1975. This area, in which the authors have both made a major experimental contribution themselves and stimulated new theoretical approaches, is considered critically and with a balanced combination of clinical and experimental evidence. Wide and important issues are covered, such as the possible reasons for a highly polymorphic system of tissue antigens and the resistance to rejection of the foetus or its malignant derivatives. Although a great deal of data is presented, it is, however, often in a form not conducive to logical exploration of the problem. For example, discussion of the question of trophoblast antigenicity was dispersed to several points in the text for no obvious reason, and at times this led to an impression of repetitiveness even when new information was imparted. At the end of this series of chapters, one was left with an element of intellectual indigestion, as though the authors had not had sufficient time to complete their synthesis of the data and analysis of the problem. We certainly do not gain a clear impression of either current opinion or the authors' opinion on the mechanism by which the mammalian foetus survives *in utero*.

Elsewhere in the volume there are some important omissions for a book

on immunoreproduction—for example, immunity to the hormones of reproduction and to ovarian tissue are not considered at all, and immunity to products of the male genital tract are discussed rather briefly. It is also evident, understandably, that the authors are ill at ease in discussing some aspects of the subject. For example, in the chapter on reproductive biology, the impression is given that uteri of all species secrete blastokinin, that the function of this protein is agreed and that 'gastrulation' occurs in the oviduct.

In spite of its deficiencies this book will be a useful source of reference for those workers actively engaged in the investigation of this problem. For others with a more peripheral interest in the subject of immunoreproduction, the volume may prove rather hard work, but will acquaint them with the varied, rather confused, but immensely exciting problems of such central importance to biology.

Martin H. Johnson

Dr Johnson is a Lecturer in Anatomy at the University of Cambridge, UK.

Conversations in culture

Neurophysiologic Studies in Tissue Culture. By Stanley M. Crain. Pp. xii+280. (Raven: New York, 1976.) \$25.50.

It is indeed a pleasure to rely on the insight of someone with almost 15 years of experience 'listening' to the conversations which nerve cells carry on in culture. Crain has put together a strong concoction of his own data as well as several others' in the field. He has added spice through description of his role as a kind of ambulance driver to transport the live cultures through hectic New York City traffic.

The monograph begins with an orientation which relates his approach to that of others in the study of brain electrical function, and includes a brief historical note on the progress through which recording from cultured cells has grown. The second chapter deals briefly with methods of growth of cells and techniques for recording action potentials with either extracellular or intracellular electrodes. The author then details the synaptic potentials which have been obtained from dorsal root and sympathetic ganglia.

The major portion of the book is then concerned with recording from explants of central nervous system tissues. In each case the recordings amply demonstrate both the prolonged slow wave effects of stimuli, the oscillatory after-discharges and the excitatory and

inhibitory post-synaptic potentials. The electrical tracings are well salted with both light and electron microscopic correlates.

The technique of dissociating brain into single cells and then studying their electrical activities in culture is only a few years old, relative to explants, but the dissociated cells have revealed that their behaviour is quite similar to that of the traditional explants. Once again, Crain discusses the peripheral nervous system as distinct from the central nervous system.

The next chapter deals with responses to various pharmacological agents, including curare, strychnine and bicuculine, as antagonists to the major transmitters, acetylcholine, glycine and GABA, respectively. There is a useful discussion of the results of serum from patients with multiple sclerosis and animals with experimental allergic encephalomyelitis. These sera have complex effects on both neurones and myelin. He then deals with the possible relationship of spontaneous fluctuations in electrical activity to the electroencephalogram. This is followed by a discussion of the important question of longer term 'trophic' interactions of neurones and the cells with which they synapse. Unfortunately, the consideration of the important 'supportive' role of glia is examined in an earlier chapter; one is therefore obliged to shuffle back to obtain the best synthesis. However, Crain certainly appreciates the relevant functional as well as developmental significance of the "neuro-glial" unit. These findings also touch the wider context of nerve-myelin trophism, as in the classical studies of M. Abercrombie and J. Z. Young.

The final chapter, and other sections throughout the book, contain the lessons of many years: beware simple morphological analysis, the need for standardised ingredients in cultures, attention to the state of the media as related to its last change and the importance of cell-cell contacts (that is, cell number). He ends on a positive note when confronted with the exceptional rise in both the number of interested workers and the collateral spin-off to other disciplines. The only obvious 'omission' is the study of the nerve cell lines, but as Crain certainly appreciates, this could be a monograph in itself. This is a masterful synthesis of not only his own extensive experience, but also the best of others, by an acknowledged founder-father of the art.

E. J. Thompson

Dr Thompson is Senior Lecturer in Neurochemistry at the Institute of Neurology, and Honorary Consultant to the National Hospital for Nervous Diseases, London, UK.